DOCUMENT RESUME

ED 269 421 TM 860 239

AUTHOR Charney, Davida H.
TITLE The Role of Elaborations in Instructional Texts: Learning to Use the Appropriate Procedure at the Appropriate Time.
INSTITUTION Carnegie-Mellon Univ., Pittsburgh, Pa. Dept. of Psychology.
SPONS AGENCY Office of Naval Research, Arlington, Va. Personnel and Training Research Programs Office.
REPORT NO TR-ONR-86-2
PUB DATE 14 Feb 86
CONTRACT N00014-84-K-0063
NOTE 158p.; Ph.D. Dissertation, Carnegie-Mellon University. The entire document contains small print.
PUB TYPE Reports - Research/Technical (143) -- Dissertations/Theses - Doctoral Dissertations (041)

EDRS PRICE MF01/PC07 Plus Postage.
DESCRIPTORS Adults; *Computer Science Education; *Decision Making Skills; Deduction; Games; Guides; Higher Education; Information Needs; *Learning Processes; *Problem Solving; *Skill Development; Textbook Content; Textbook Preparation; *Textbooks
IDENTIFIERS Box World (Computer Game)

ABSTRACT

The content of textbooks teaching the use of a computer was studied. An important component of skill learning, choosing the right procedure at the right time, is difficult in computer education because the connection between real-world goals and the generic procedures described in a manual is often obscure. Furthermore, it is difficult to tell which of several similar procedures is best for a particular situation. Various types of advice about choosing the correct procedure were investigated. College students and staff members read one of four versions of a manual for a computer game called "Box World." A Dandylion computer was used. Three versions contained advice about when to use particular game procedures. The advice was either stated as a simple verbal rule or elaborated with a concept example or a task example. The fourth version contained no advice. Subjects then performed three tasks: recall; recognition of correct versus incorrect applications of the advice; and decision making and problem solving. Subjects who received advice followed it to select a procedure; however, so did subjects receiving no advice. The examples helped subjects identify correct applications of the advice, but did not increase adherence to the advice during problem solving. (Author/GDC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED269421

TM

THE ROLE OF ELABORATIONS IN INSTRUCTIONAL TEXTS:

Learning to Use the Appropriate Procedure at the Appropriate Time

Davida H. Charney

Carnegie-Mellon University

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

Technical Report No. ONR-86-2
February 14, 1986

The work reported here was sponsored by the Office of Naval Research, Contract No. N00014-84-K-0063, Contract Authority Identification Number NR667-529.

TM 860 239

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) Technical Report No. ONR-86-2	5. MONITORING ORGANIZATION REPORT NUMBER(S) Technical Report No. ONR-86-2
6a. NAME OF PERFORMING ORGANIZATION Davida Charney 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS (City, State, and ZIP Code) Department of Psychology Carnegie-Mellon University Pittsburgh, PA 15213	7b. ADDRESS (City, State, and ZIP Code)
8a. NAME OF FUNDING/SPONSORING ORGANIZATION Office of Naval Research 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER N00014-84-K-0063
8c. ADDRESS (City, State, and ZIP Code) Personnel and Training Research Program Arlington, VA 22217	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. 61153N; PROJECT NO. RR04206; TASK NO. RR04206; WORK UNIT ACCESSION NO NR667-529

11. TITLE (Include Security Classification)
The Role of Elaborations in Instructional Texts: Learning to Use the Appropriate Procedure at the Appropriate Time (unclassified)

12. PERSONAL AUTHOR(S)
Charney, Davida

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Technical Report	FROM 1/85 TO 10/85	86-2-14	163

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Elaborations, computer documentation, human-computer interaction, skill learning, text comprehension, examples, selection strategies

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

This research investigates the kinds of information that should be included in instructional texts that teach skills (such as manuals or textbooks). It focuses on an important subcomponent of skill learning: choosing the right procedure at the right time. Learning to choose the right procedure is difficult in skills such as using a computer because the connection between real-world goals and the generic procedures described in a manual is often obscure. Furthermore, when several procedures have similar functions, it is difficult to tell which one is best for a particular situation. In order to facilitate the decision process, instructional texts may include advice about when to use particular procedures and may illustrate the advice with examples. The research reported here investigated the effect of various forms of advice on learners' strategies for choosing a procedure. Subjects read one of four versions of a manual for a computer game called Box-World. Three versions contained advice about when to use particular game procedures; the advice was either stated as a simple verbal rule or elaborated with one of two types of examples. The fourth version contained no advice. Subjects then

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL Michael Shafto	22b. TELEPHONE (Include Area Code) / 22c. OFFICE SYMBOL

DD FORM 1473, 84 MAR — 83 APR edition may be used until exhausted. All other editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE
unclassified

19. (cont.) performed three tasks: a recall task, a recognition task (i.e., discriminating between correct and incorrect applications of the advice), and a decision task (i.e., solving problems for which the advice was relevant). Subjects who saw certain forms of the advice were equally able to solve the problems correctly. Examples helped subjects identify correct applications of the advice, but did not increase adherence to the advice during problem solving. The results suggest that advice does not require elaboration with examples, but further research is required on how subjects interpret advice and decide to apply it.

CARNEGIE-MELLON UNIVERSITY
College of Humanities and Social Sciences

DISSERTATION

THE ROLE OF ELABORATIONS IN INSTRUCTIONAL TEXTS:

Learning to Use the Appropriate Procedure at the Appropriate Time

Submitted in Partial Fulfillment of
the Requirements for the Degree of
Doctor of Philosophy in Rhetoric

by

Davida Charney
October 11, 1985

Table of Contents

ABSTRACT 1

ACKNOWLEDGMENTS 3

1. INTRODUCTION: DESIGNING EFFECTIVE TEXTS 5

   1.1. A Note on Methodology 7
   1.2. An Overview of the Dissertation 9

2. THE ISSUES OF CONTENT AND ELABORATION IN INSTRUCTIONAL TEXTS 11

   2.1. The Traditional View on Content 11
   2.2. Research on the Role of Elaborations in Skill Learning 13
   2.3. The Effect of Elaborations on Skill Performance 17
   2.4. Controlling What Types of Information Receive Elaboration 22
      2.4.1. The Implications of the Results 24

3. THE ROLE OF SELECTION INFORMATION IN SKILL LEARNING 29

   3.1. Selection Information and Task Characteristics 29
   3.2. What It Takes to Follow Advice 34
   3.3. Finding the Most Effective Form of Advice: the role of examples 36
      3.3.1. The Benefit of Examples for Concept and Skill Learning 36
      3.3.2. How Examples Aid Learning 38
      3.3.3. Conceptual Examples and Task Examples 43
   3.4. A Preview of the Experiment 44

4. METHOD 49

   4.1. Overview of the Box-World Game 49
   4.2. Design 52
      4.2.1. Efficiency Task 53
      4.2.2. Advice Recognition Task 54
   4.3. Materials 55
      4.3.1. Instructional Materials for Box-World 55
      4.3.2. Trials for the Box-World Game 58
      4.3.3. The Move Isomorph 61
   4.4. Apparatus 62
   4.5. Subjects 62
   4.6. Procedure 62
      4.6.1. Efficiency Task 63
      4.6.2. Recall Task 65
      4.6.3. Advice Recognition Task 65

5. RESULTS AND DISCUSSION, PART I: THE EFFICIENCY TASK 69

   5.1. Overview 69
      5.1.1. A Note on the Number of Subjects per Task 70

5.2. The Utilization and Utility of the Advice: the Efficiency Task 71
5.3. Did Subjects Follow the Advice? 73
5.4. Was the Advice Helpful? 74
5.5. Did Task Difficulty Affect Reliance on Advice? 80
5.6. Did the Form of the Advice Affect Performance? 81
5.6.1. The Effect of the Task Example 81
5.6.2. The Effect of the Concept Example 84
5.7. The Effect of the Move Isomorph 86
6. RESULTS AND DISCUSSION, PART II 95
6.1. Recall Task 95
6.1.1. Recall of Advice 95
6.1.2. Recall of Examples 98
6.1.3. Retrospective Report on Solution Strategy 100
6.2. Advice Recognition Task 100
6.2.1. Accuracy of Decisions 101
6.2.2. Reaction Time 104
6.3. Replication effects 106
7. CONCLUSION 109
7.1. The Role of Elaborations in Instructional Texts 109
7.2. Goals for Future Research 112
Appendix A. BOX-WORLD MANUAL 115
Appendix B. BOX-WORLD TASKS 121
Appendix C. BOX-WORLD TASKS – MOVE ISOMORPHS 131
Appendix D. DISTRIBUTION OF SUBJECTS AMONG CONDITIONS 141
Appendix E. SAMPLE RECALL RESPONSES 143
REFERENCES 145

Tables and Figures

Table 2-1	Mean Steps per Subject for Five Kinds of Actions as a Function of Version of Manual	20
Figure 4-1	Box-World Situation with Goal	50
Figure 4-2	Box-World Situation with Action	51
Figure 4-3	Three Forms of Advice for the Change Command	56
Figure 4-4	Three Forms of Advice for the Move Command	57
Figure 4-5	Instructions for the Efficiency Task	64
Figure 4-6	Questions in the Cued-Recall Task	67
Figure 4-7	Instructions for the Advice Recognition Task	68
Table 5-1	EFFICIENCY TASK: Percent Correct Decisions as a Function of Instruction, Appropriateness and Difficulty (Change Isomorph)	72
Table 5-2	EFFICIENCY TASK: Encoding and Planning Time (in seconds), as a function of Instruction, Appropropriateness and Difficulty (Change Isomorph)	76
Table 5-3	EFFICIENCY TASK: Decision Time (in seconds), as a function of Instruction, Appropriateness and Difficulty (Change Isomorph)	78
Table 5-4	EFFICIENCY TASK: Percent Correct Decisions as a Function of Instruction, Appropriateness and Difficulty (Move Isomorph)	87
Table 5-5	EFFICIENCY TASK: Encoding and Planning Time (in seconds), as a function of Instruction, Appropropriateness and Difficulty (Move Isomorph)	89
Table 5-6	EFFICIENCY TASK: Decision Time (in seconds), as a function of Instruction, Appropriateness and Difficulty (Move Isomorph)	90

Table 6-1	RECALL OF ADVICE. Mean Recall Scores as a function of Instruction and Isomorph	99
Table 6-2	ADVICE RECOGNITION TASK. Percent Correct Decisions as a function of Isomorph, Instruction, Match and Difficulty	102
Table 6-3	ADVICE RECOGNITION TASK. Overall Reaction Times (in seconds), presented as a function of Isomorph, Instruction, Match and Difficulty	105

ABSTRACT

This research investigates the kinds of information that should be included in instructional texts that teach skills (such as manuals or textbooks). It focuses on an important subcomponent of skill learning: choosing the right procedure at the right time. Learning to choose the right procedure is difficult in skills such as using a computer because the connection between real-world goals and the generic procedures described in a manual is often obscure. Furthermore, when several procedures have similar functions, it is difficult to tell which one is best for a particular situation. In order to facilitate the decision process, instructional texts may include advice about when to use particular procedures and may illustrate the advice with examples. The research reported here investigated the effect of various forms of advice on learners' strategies for choosing a procedure.

Subjects read one of four versions of a manual for a computer game called Box-World. Three versions contained advice about when to use particular game procedures; the advice was either stated as a simple verbal rule or elaborated with one of two types of examples. The fourth version contained no advice. Subjects then performed three tasks: a recall task, a recognition task (i.e., discriminating between correct and incorrect applications of the advice), and a decision task (i.e., solving problems for which the advice was relevant). Subjects who saw certain forms of the advice followed it to select a procedure. However, subjects who saw no advice were equally able to solve the problems correctly. Examples helped subjects identify correct applications of the advice, but did not increase adherence to the advice during problem solving. The results suggest that advice does not require elaboration with examples, but further research is required on how subjects interpret advice and decide to apply it.

ACKNOWLEDGMENTS

Before acknowledging any other intellectual debts, I would like to thank my committee. I feel very fortunate to have worked with Lynne Reder, who has been most generous with her time, her resources and, most valuable, her expertise. Lynne taught me most of what I know about conducting experimental research and her work on elaborations and memory underpins a lot of my thinking. In the course of this project, as on numerous previous occasions, I found Dick Hayes to be a problem solver extraordinaire. He was unfailingly encouraging and continually provided thoughtful, creative and insightful advice. Tom Duffy not only kept me alert to issues of experimental design and lesher statistical analyses, he also raised my consciousness to alternative psychological paradigms. Dave Kaufer helped me unearth hidden assumptions and make explicit the connections between my work and the goals of our discipline. He followed my arguments and plans, anticipated where they were heading, and often steered them down better if not more easily travelled paths. Of course, any blunders that somehow slipped past all these eagle eyes are entirely my responsibility.

This work represents part of a research contract supported by the Office of Naval Research, Contract No. N00014-84-K-0083. I am very grateful to ONR for funding the contract and to Lynne Reder for allowing me the time and resources to carry out this part of the project. I also owe a special debt of gratitude to Bob Allison who developed the Box-World programming even though his official connection to the ONR project had ended.

While enviously paging through completed dissertations in my early graduate days, I noticed that the most heartfelt thanks often went to whoever the author had managed to rope in to type (and retype) the manuscript. My good fortune is that thanks to the outstanding computer facilities at CMU, I didn't need a typist. This leaves me free to express my most heartfelt thanks to those people in the English and Psychology departments at CMU who made this a rich, stimulating place to work.

Everyone is bound to bear with patience
the results of his own example.
—Phaedrus

Chapter 1
INTRODUCTION: DESIGNING EFFECTIVE TEXTS

Helping writers to produce more effective texts is an important goal of both writing pedagogy and research on document design. The sticking point in the previous sentence is, of course, the phrase *more effective*. What makes one text more effective than another? Rhetoricians know that the aim of a discourse may be to instruct, delight, move or persuade the reader. In some sense, then, a more effective text simply carries the author's intention further; the persuasion is deeper or the delight is greater. This way of looking at effectiveness presupposes two things: first, that you can isolate exactly what the text is supposed to accomplish, and second, that you can measure the effect of the text on the reader.

For some types of text, it is not at all easy to decide what effect the text is supposed to have. For example, literary texts might not have any determinate, "intended" effects, so the question of effectiveness may be irrelevant or it may depend entirely on which theory of literature the reader adopts. Even when the discourse has an identifiable aim, the effect on the reader may be difficult to discern or measure. Argument, for example, aims at changing the reader's beliefs. It is difficult, however, to measure the extent to which a person has been persuaded by a text, since a change in beliefs may not always manifest itself in the reader's behavior in a predictable way, or in measurable degrees.

Functional documents, especially instructional texts, have two qualities that make *them particularly accessible to research on effectiveness. First, functional documents* have at least some overt, determinate goals. For example, the instructions for the 1040 tax form are intended to help people fill out their tax return correctly. A computer user's manual is intended to help people learn to operate the computer. So, while a functional document can certainly have hidden agenda, the primary goals of the document are usually quite explicit. Second, functional documents have observable and measurable effects on their readers. After reading the text, people should be able to do something that they simply couldn't do before (or couldn't do as well).[1] Therefore, it is relatively easy to decide whether a functional text is effective, or which of two functional texts is more effective. An effective text helps people learn to do something. A more effective text helps people learn to do that something more easily and quickly, with fewer mistakes or problems. Because functional texts have these qualities, it is possible to explore the features that distinguish effective texts from ineffective ones.

My long-term goal is to develop principles for deciding what to say and how much to say in an instructional text in order to teach a skill most effectively. For the purpose of this work, I will limit the term *instructional text* to include such texts as geometry textbooks and computer user's manuals, which emphasize skill performance and problem solving, but exclude texts such as history books and newspaper articles, which place more emphasis on the acquisition and synthesis of information. An important assumption of the work is that principles for designing

---

[1]Obviously, acquiring a skill in any true sense requires a great deal of first-hand experience. However, learning the relevant concepts, principles and procedures in verbal form is a common and almost indispensable first step, as Anderson (1983) points out.

instructional texts must be grounded in a cognitive model of what information people need to perform the skill and how they use this knowledge to construct and refine new mental procedures. This work therefore draws heavily on cognitive psychology for theories of text processing and skill learning and for experimental methods of evaluating skill performance under different learning conditions (that is, how well readers can perform new skills after studying different instructional texts).

1.1. A Note on Methodology

In recent years, there has been a movement away from analyzing texts as isolated artifacts and toward viewing texts as part of a complex relationship among the writer, the reader, the external situation being written about and the text that describes it (e.g., Young, Becker & Pike, 1970; Kinneavy, 1971). This movement has influenced the way in which texts are evaluated; more and more, the intended reader is being included in the process. One example of this trend is the growing popularity of holistic scoring for evaluating writing ability. Many people believe that holistic scoring is more valid than so-called analytic measures because it seeks a natural response to a text from a reader (Charney, 1984a). Similarly, more and more writers of functional documents are taking the sensible step of seeing what a typical person actually learns from reading their documents.

User-testing (or the "user-edit," as Atlas, 1980, termed it) has been an important development for both researchers and practitioners of document design. By observing typical members of the target audience using a document to perform a typical task, writers gain valuable information about where the document works and where it doesn't. With the addition of thinking-aloud protocols (as described, for example, in Swaney, Janik, Bond & Hayes, 1981; Sullivan and Flower, forthcoming; and Charney,

1984b) user-testing can also suggest causes for the problems and point to possible solutions. Repeated cycles of testing and revision can result in highly effective documents.

The primary goal of user-testing is to make individual documents more effective. The primary goal of this dissertation, however, is to develop general principles for creating effective instructional texts. Consequently, this dissertation sets the user-testing approach within an experimental framework. By systematically controlling the features in various versions of a text and comparing what readers learn from the different versions, we can gain a better understanding of which features are most important and why they have the effect they do.

This dissertation is related both to cognitive psychology and to the work of instructional design theorists such as Merrill (1983) and Reigeluth & Stein (1983). Instructional designers have developed a family of theories about the best forms of instruction. Like cognitive psychologists working in the information processing paradigm, instructional designers begin with a detailed analysis of the task that learners are to perform. The objectives of the two approaches differ, however. Information processing models aim at describing how people naturally go about solving problems in order to illuminate human mental resources and capacities. Instructional design theories aim at guiding learners to solve problems in the optimal way. So, after setting well-defined performance objectives, an instructional designer selects, elaborates and organizes information in ways prescribed by his or her theory to create instructional materials for achieving those objectives. In this dissertation, the interests of composition teachers, instructional designers and cognitive psychologists overlap: in order to teach people to write effective functional texts, we can benefit from research on how people learn as well as research on optimal forms of instruction.

1.2. An Overview of the Dissertation

The research in document design that has adopted an experimental methodology has mainly focused on two sorts of goals: making the information in the text easier to find and making the information easier to understand. So, a great deal of document design work has centered on problems with technical language or on concerns like the role of tables of contents and indices, or page layout and typography (e.g., Wright, 1977, 1978, 1983; Duffy & Kabance, 1982; Kieras, 1985; and research reviewed in Felker, 1980). An important issue that has not received enough attention is the issue of information content. What kinds of information should an instructional text include? How much should be explained and what form should the explanations or elaborations take?

This dissertation is part of a larger research project that focuses on information content and, in particular, on the role of elaborations in skill learning (Reder, Charney & Morgan, in press). Chapter 2 begins by presenting the traditional view of information content. It then reviews research that points to a more complex view of the role of elaborations for learning a skill from an instructional text. My particular focus is on an important subcomponent of learning a skill: learning to choose the right procedure at the right time. Chapter 3 introduces the problem of selecting appropriate procedures, why this is an important problem and how advice and examples in an instructional text might help people learn a selection strategy. Chapter 4 describes the experimental method of the research study on advice and examples. The results of the study are presented and discussed in Chapters 5 and 6. Finally, Chapter 7 presents some concluding remarks on the implications of the research for document design, including goals for future research.

Chapter 2
THE ISSUES OF CONTENT AND ELABORATION IN INSTRUCTIONAL TEXTS

In this chapter, I will sketch the traditional view on the content of instructional texts, the view that all points should be explained in complete detail. After presenting some reasons for doubting this traditional wisdom, I will briefly review some experimental research that leads to a more complex view: that elaborations are only needed for certain kinds of information that relate to a specific component of skill learning.

2.1. The Traditional View on Content

The traditional view on content is that an instructional text for novices should assume little or no prior knowledge. The text should be as complete and as explicit as possible with detailed explanations of every relevant point. This point of view is held by writers of both textbooks and instructional manuals. Tausworthe (1979), for example, outlines several levels of detail for documenting computer software. The highest level of detail is called for in what he labels "Class A documentation," which he describes as follows:

> Class A documentation is the most detailed; it contains specific definitions and detailed descriptions of every significant factor or item within the software specification..... This level of detail probably finds its most applicability in user manuals, and rightly so: The writer of a user manual is generally unavailable for consultation, so the user needs the extra detail. [Tausworthe (1979), pp. 158-159.]

This is a good statement of the traditional wisdom. But, as it often happens, there is good reason to doubt the traditional wisdom. For one thing, real computer users apparently don't like to read long, detailed instructional manuals. They'd rather ask someone to show them what to do or try to figure it out themselves (Scharer, 1983). Furthermore, Pirolli and Anderson (1985) have observed that even when the learners do read the text, they draw on relatively little of this information during task performance. So from this point of view, providing complete and detailed instruction seems of little practical value to the learner.

A second reason to doubt the traditional wisdom is provided by research in the fact-learning domain. In a series of ten experiments, Reder and Anderson found that subjects who studied a full length textbook chapter performed worse on comprehension and recall tests than subjects who had studied summaries that were one-fifth as long (Reder & Anderson, 1980; Reder & Anderson, 1982; Allwood, Wikstrom & Reder, 1982; Reder, 1982).

Reder, Charney & Morgan (1984, and in press) explain these rather surprising results as a combination of two phenomena, one involving encoding and the other retrieval. The encoding phenomenon is the well-known Total Time Law (Bugelski, 1962; Cooper & Pantle, 1967): The more time and attention a person spends studying a particular fact, the better he or she will learn it. So, reading elaborations in the text diverts time and attention away from the main points. On the other hand, studying a summary allows the reader to devote full attention to exactly those points that need to be remembered. The encoding phenonmenon is not the whole story, however. In one study, Reder and Anderson (1982) equated the total time subjects spend studying the main points. The subjects in the elaborated group spent extra time studying the elaborations. Presumably, the encodings for the main points

in both groups were equally strong because subjects in both groups spent the same amount of time studying them, but there was still a significant advantage for the summary group. So, the handicap of the elaborations seems also to involve a second phenomenon, retrieval. When the learner tries to retrieve the main points of a text from memory, the elaborations cause interference (Anderson, 1974; Reder & Anderson, 1980).

The findings from the fact-learning domain suggest that elaborations hurt performance when the learner's goal is to understand and remember the main points of a text. But what happens when the goal is to apply the facts, to use them to solve problems, as in skill learning? We know that learners don't like having to rely on computer manuals, but it's an open question why computer manuals and other instructional texts are so unsatisfactory. It could be that such texts fail to live up to the traditional wisdom: that is, they assume too much prior knowledge and so don't contain enough information. Or it could be that they contain too much information that distracts attention away from the essential facts; that is, the advantage of summaries might apply even in the skill learning domain.

2.2. Research on the Role of Elaborations in Skill Learning

To test whether elaborations help or hurt skill learning, we conducted a study on the user's manual for the Disk Operating System (DOS) of the IBM Personal Computer (for a more complete report of this study, see Reder, Charney & Morgan, 1984 and in press). We rewrote part of the original IBM manual (Anonymous, 1983) to include the clearest and most relevant elaborations we could think of, such as detailed explanations, analogies, metastatements, examples of commands, and so on. This "elaborated" version of the manual came to about 40 pages. Then we created

a second, "unelaborated" version of the manual by simply deleting all of the elaborations. This manual came to about 12 pages.

Eighty novice computer users were given up to an hour to read one version of the manual and were told that the manual would not be available to them after the reading period was over. After the subjects read the manual, they were asked to carry out a set of four ordinary tasks on the computer: renaming files, creating subdirectories, copying and deleting files, and so on. As the subjects worked, the computer kept a record of every command they typed and the time at which it was entered. The measures of how well subjects performed were whether they were able to do the tasks and how efficiently they worked (e.g., how much time they took and how many commands they issued).

Half of the subjects, the "Before" group, were given advance information about the tasks they were going to perform, before they read the manual. The "After" group saw the task instructions only after they had finished reading the manual. This manipulation was included in order to simulate two common learning situations. Sometimes learners have explicit goals in mind and turn to instructional materials specifically to find information relevant to those goals. At other times, people come to learn a new skill with only a general idea of how they will make use of what they learn. We speculated that elaborations provided in the text might be less valuable to the Before group, who could presumably generate their own task-specific elaborations while reading.

We found that the version of the manual did not affect the number of tasks subjects could complete, suggesting subjects could learn enough from any of the manuals to complete most of the tasks, providing that the time for working on the

tasks was not restricted. However, we did find differences in how efficiently subjects worked. Subjects performed better if elaborations were available, either elaborations in the text or elaborations that they generated themselves based on advance knowledge of the tasks; they were faster at performing the tasks and issued fewer commands. However, having either one of these sources of elaborations was sufficient: having both elaborations in the text and advance knowledge did not boost performance above one of these sources alone.

For some measures, we found that the shorter, unelaborated manual worked well for subjects who had advance knowledge of the tasks. However, the results in general support the idea that manuals should contain elaboration. Realistically, we can't assume that all learners will come in with such clearly defined goals as the Before group subjects. In fact, the worst performance always came from subjects who had no elaborations available, i.e., subjects who studied the unelaborated manual without advance knowledge of the tasks. On balance, these learners were impeded more by the under-elaborated texts than the more goal-directed learners were by the over-elaborated version. At this point, then, the best strategy would seem to be to play it safe and provide elaborated instruction to all learners.

So these results seem to support the traditional wisdom of giving novices complete explanations of all the relevant points. Unfortunately, we can't lay the question to rest here, since Carroll (1985) did a fairly similar study that produced conflicting results.

Carroll and others have adopted a new philosophy toward instructional manuals, the so-called Minimalist philosophy. Designers of so-called "minimalist training materials" proceed on the assumption that instructional materials should actively

encourage discovery learning by providing as little prose as possible. Carroll put this principle into practice in a tutorial manual for a commercial word processing system (the IBM Displaywriter System) and produced a revised manual that was one-fourth the length of the original. Carroll's principles for shortening the manual included two major steps. First, he cut out everything he considered irrelevant to the task at hand.

> ...eliminating all repetition, all summaries, reviews, and practice exercises, the index, and the troubleshooting appendix. ...All material not related to doing office work was eliminated or radically cut down (the welcome to word processing overview, descriptions of the system status line, details on the system components..., etc.). [p.5]

Carroll's second step was to take what was left, the relevant information, and delete parts that he believed learners would be able to figure out on their own.

> Procedural details were deliberately specified incompletely to encourage learners to become more exploratory, and therefore, we hoped, more highly motivated and involved in the learning activity (e.g. the function of the cursor step-keys was introduced with an invitation to 'Try them and see.') [Carroll, p.6]

Carroll's manual was tutorial in the sense that readers were expected to try things out as they read about them. After giving some subjects the Minimal manual and other subjects the original manual, Carroll found two things: first, subjects who worked through the minimal manual learned the basic information more quickly, and second, when these subjects went on to study advanced techniques, they learned more of these techniques more quickly than subjects who worked with the original materials.

Carroll admits that this effort at designing minimalist materials was exploratory. For example, as a result of preliminary testing, he found he had to add in some explanatory sections as well as some procedures that subjects actually couldn't figure out on their own. Furthermore, Carroll made other changes to the manual in

addition to making it shorter: he clarified the terminology and organized the discussion around typical situations for users. It is therefore uncertain how much of the superiority of the minimalist version is due to length and how much to these other changes. However, in the main Carroll's findings support the minimalist position: having less to read led to equivalent or better learning at a faster overall rate.

Neither Carroll's study nor ours allows us to generalize about whether or not elaborations should be included in a text. We have no systematic basis for contrasting the kind of elaborations in our manual to those in the original Displaywriter manual, so we don't know what we did right to get better performance and what the writers of Carroll's elaborated version did wrong. Furthermore, we don't know exactly what our elaborations did to improve subject performance, so we can't predict what kind of elaborations are worthwhile. The work I will describe now pursues both of these questions: what types of elaborations are necessary and how they affect the user's behavior at task.

2.3. The Effect of Elaborations on Skill Performance

The subjects in our experiment who studied the elaborated manual completed the tasks in fewer steps than subjects who saw no elaborations. There are two major ways in which the elaborations might have helped subjects perform the tasks more efficiently. First, elaborations may have helped subjects construct more efficient plans for accomplishing a task. Some elaborations in our manual gave advice about when to use short-cuts (such as wildcard characters). If these elaborations helped subjects remember to use the short-cuts at the appropriate times, then subjects who read the elaborated manual would be able to complete the tasks with fewer commands. The second area in which elaborations may have helped is to generate syntactically

correct computer commands. The elaborated manual contained many examples of syntactically correct commands and detailed explanations of what the notation meant. These elaborations may have helped subjects remember the names and the syntax of the commands and formulate syntactically correct commands more easily.[2]

We investigated these alternatives by analyzing the kinds of commands subjects issued and the kinds of mistakes they made. These data were available from the records or "on-line protocols" of the subjects' interactions with the computer. The commands in the protocols were sorted into five categories:[3]

- Productive moves: syntactically correct commands that carry out a "target action" or that enable one.
- Verification moves: commands that check whether a previous command had the desired affect.
- Execution errors: commands that contain one or more syntactic errors.
- Goal specification errors: wrong command issued or failure to perform a prerequisite action. (The subject may have had some misconception about current state of the computer or the capabilities of a command.)
- Recovery moves: commands to gain information after an error or to undo undesired effects.

It is important to note that recovery moves are not simple corrections of syntactic errors; recovery moves are efforts at problem solving to diagnose syntactic errors (such as rearranging the elements of the command line), as well as efforts to figure

---

[2]Examples of syntactically correct commands may help in other ways than illustrating the command syntax. Seeing a variety of examples depicting different applications of a command might also help people learn more about the function of the command or when or in what circumstances to apply it. This function of examples was not exploited in our manual, since most commands were illustrated with just one example.

[3]In all, protocols for 20 subjects were analyzed; half had studied the Elaborated version of the manual and half the Unelaborated version. All of the subjects were in the After condition (i.e., none had had advance knowledge of the tasks).

out what effect a command had and to undo any undesired consequences (such as deleting files that were copied to the wrong location). Any syntactically correct command that achieved a target action was classified as a productive move, regardless of whether it was preceded by an incorrect attempt.

Other researchers of human-computer interactions have used schemes to analyze the errors people make while performing tasks on the computer, for example Riley & O'Malley (1984) and Sebrechts, Galambos, Wagner, Black, Deck and Wilder (forthcoming). These coding systems differ in some important respects; for example, only our system examines the relative distribution of productive commands and different types of errors. However, all three systems distinguish between goal setting problems and problems of execution.

Table 2-1 shows the distribution of the five types of commands for subjects who saw the elaborated manual and subjects who saw the unelaborated manual. The number of steps for productive and verification moves were essentially the same for the two groups. If the elaborations had helped subjects invent more efficient solution strategies, then we would have expected the elaborated group to have needed fewer productive moves. But this was not the case.

The difference in behavior between the two groups appears in the final three categories of commands. Subjects in the Elaborated condition issued less than half as many commands that contained syntactic errors as subjects in the Unelaborated condition, $t(18) = 2.6$, $p < .05$. About the same ratio held for goal-specification errors and recovery moves, but only the contrast for the recovery moves was significant, $t(18) = 2.3$, $p < .05$. So two things that the elaborations seemed to help subjects do were to learn to generate correct commands and figure out how to fix bad ones.

TABLE 2-1

Mean Steps Per Subject for Five Kinds of Actions
as a Function of Version of Manual

	Elaborated Manual	Unelaborated Manual
Productive Moves	27.7 steps	33.7 steps
Verification Moves	11.0	12.3
Execution Errors	9.5	20.2*
Goal Specif. Errors	7.3	13.5
Recovery Moves	11.3	20.8*

*Contrast is significant at the .05 level

Reder, Charney and Morgan (in press) sets these results into a bigger picture of what people have to learn in order to use a new set of procedures, like those in a computer manual. We conceive of good skill performance on a novel task as requiring mastery of three sub-components:

1. Appreciating the meaning of novel concepts and procedures. For example, what is a subdirectory, what does it mean to copy a file?

2. Remembering the procedures and understanding how to execute them in a specific situation. For our task domain, this involves remembering the syntactic rules and using them to generate correct commands.

3. Remembering to use the most appropriate procedure for the situation at hand. When there is more than one way to do what you want, you need some selection rule or strategy.

Granted that the protocol analysis was *post hoc* and that this evidence is not reinforced by thinking-aloud protocol data, the results of the analysis are still quite suggestive. The advantage of the elaborated group seems to be in the second subcomponent: knowing how to formulate commands correctly, recognizing what was wrong in syntactically incorrect commands and figuring out how the computer interpreted the incorrect commands.

How the information in the manual can concern any of these subcomponents: Information about what commands are available and what they do, information about how to issue correct commands and advice on when to use one command rather than another. But some of the elaborations in the manual seem much more relevant than others to learning how to formulate correct commands. Presumably, the most relevant elaborations would be those that relate directly to the command notation and those that give examples of correct commands. If the only advantage of the elaborated version was in helping people learn the command syntax, then perhaps only the elaborations relevant to this topic were necessary in the manual.

2.4. Controlling What Types of Information Receive Elaboration

In order to test whether the syntactic elaborations were the source of the difference in subjects' performance, we conducted another study in which we systematically controlled the availability of the elaborations by type (again reported more fully in Reder, Charney & Morgan, in press). We started with the elaborated manual we had used in the previous experiments. Then we classified the elaborations in the manual according to whether they concerned "conceptual" or "syntactic" information. Elaborations were classified as "conceptual" if they concerned basic concepts, such as the purpose of a command or when to use it. Elaborations were classified as "syntactic" if they concerned how to issue commands correctly (e.g., examples of commands, details about notation conventions, etc.). By crossing these factors, we came up with four versions of the manual; one contained both conceptual and syntactic elaborations, one contained just the conceptual elaborations, one just the syntactic ones and one contained neither type.

The purpose of the classification was to separate out information about syntax from other types of information. Two caveats are in order. First, although we have classed "functional" information about the purpose of a command together with "selection" information about when to use a command, this should not necessarily be taken as a theoretically significant grouping. In fact, the research to be described in this dissertation will attempt to sort out the role of selection information. Second, we have classed examples of syntactically correct commands as syntax elaborations. As noted above, such examples may have other benefits relating to learning the function or appropriate application of commands. To the extent that subjects need additional elaboration for these functions, we would expect the "conceptual" elaborations to improve performance. If the syntax examples also contribute heavily to these

functions (as opposed to the syntax function we have postulated), then we would expect to see similar benefits from both types of elaboration.

In this experiment, half of our subjects were experienced computer users and half were novices. We expected these groups to have complementary needs as far as the conceptual elaborations were concerned. Experienced computer users already understand the general concepts behind computer systems. They need to know what options are available on a new system, but they might be distracted and bored by elaborations on concepts they already understand. Novices however lack a clear conception of what a computer operating system can do. We expected them to benefit from fuller explanations. On the other hand, we expected experienced computer users and novices alike to benefit from elaborations on the command syntax. Experienced computer users are probably better at parsing the abstract syntactic rules that are found in most computer manuals, but both experienced computer users and novices should be better at formulating correct commands themselves if they have seen examples of correct commands. So both groups of subjects should benefit from elaborations about how to issue syntactically correct commands.

Not surprisingly, we found that the experienced computer users performed much better than the novices on every measure. However, we didn't find that experts and novices needed elaborations on different types of information. Instead, both groups benefited from the syntax elaborations; When the syntax elaborations were available, subjects worked significantly more quickly and issued significantly fewer commands than when they were absent. In no case did the conceptual elaborations seem to have any effect whatsoever.

These results provide striking evidence that users benefit from extra information on how to execute procedures: both experienced and novice computer users performed better with such elaborations than without them. The finding that the conceptual elaborations were useless even for the novice computer users was a bit surprising. However, it seems premature to conclude that we never need to elaborate on conceptual information in instructional texts.

2.4.1. The Implications of the Results

My goal in this dissertation is to consider further the question of the conceptual elaborations, in particular those that concern "selection" information: when to use particular commands or how to choose between similar commands. There are several possible explanations for why elaborations on selection information failed to improve performance while syntax elaborations succeeded.[4] Let's consider three plausible explanations. First, it is possible that both types of elaborations really are useful, but the selection information happened to be irrelevant to the particular tasks in this experiment (i.e., deciding among the plausible commands for each task may have been trivially simple or the selection of one command over another may not have drastically improved or impaired efficiency). If so, then it should be easy to show that elaborations on selection information improve performance whenever tasks require careful choices among commands.

A second explanation shifts the blame from the tasks to the elaborations themselves. Suppose that the selection information really was relevant to good performance on these tasks (i.e., the choice between commands was hard or would

---

[4]The arguments developed in this section focus on selection information, but corresponding arguments can be developed for "functional" information as well.

have affected efficiency), but subjects couldn't figure out how to take advantage of this information and the particular elaborations in the manual happened not to be the most effective kind. As a result, subjects failed to *make* good use of the selection information, whether elaborations were available or not. However, if the manual had contained better elaborations, then subjects would have been able to exploit the selection information and their performance would have improved. In this case, the trick is to find the right kind of elaboration.

The third and most interesting possibility is that the results of this study should be taken at face value: there may be different types of information, only some of which need extensive elaboration in instructional manuals. Information about how to generate syntactically correct computer commands would then seem to be a type that benefits from elaboration, but selection information is not. In this case, we should never find improvement for subjects who read selection elaborations, no matter what the elaborations look like, even when the experimental tasks require careful selection among commands.

Notice that these explanations posit three features to control whether elaborations have a chance to improve performance: the tasks must be ones for which the information is relevant, the elaborations must be the right kind and the information must be a type that benefits from elaboration. Each explanation assumes that a different one of these three features was missing for the selection information in our experiment. It is interesting and important to find out which explanation is correct because they have different implications for writing effective manuals. It seems intuitively obvious that writers will always have to consider the tasks their readers will

perform and that they must always choose elaborations carefully.[5] However, if the third explanation is correct, then there is an additional feature, information type, that may help writers decide whether or not to elaborate some point in a manual.

Let's begin by following out the implications of the first and second explanations. Suppose that for any given kind of information, we can always find a task for which performance is improved when the right elaborations are available. Then, in order to decide whether or not to elaborate a point, writers never need to consider what kind of information the point is, but only the relative abundance of tasks for which that information is relevant. For example, suppose that there are tasks on the IBM-PC for which selecting the most appropriate command is difficult and making the right choice greatly improves the efficiency of the solution. If these "selection-critical" tasks or situations are fairly common, then selection information should be included in the manual to help the computer user pick the best command. On the other hand, these situations might be relatively rare. That is, the tasks we chose for our experiment may have been highly representative of the tasks most people perform on the IBM-PC most of the time. Then it might be safe to omit the selection elaborations and include only syntax elaborations, as we did in one of our manuals. In short, if we can assume that any type of information can benefit from elaboration, then writers only need to worry about the distribution of tasks for which the information is relevant and finding the right elaborations.[6]

---

[5]Holland, Rose, Dean & Dory (1985) provide some experimental evidence that even for the "syntax" type of information (i.e., how to execute a procedure), the level of specification must be carefully chosen; too fine a degree of detail only confuses the learner.

[6]Under this analysis, the reason why syntax elaborations are so helpful is that no matter what task you may wish to use a command for, you will always need to know how to generate the command line correctly.

On the other hand, if the third explanation is correct, then it's not sufficient to consider the tasks and the quality of the elaborations. Writers must also consider the type of information. It would be very helpful to know in advance that syntax information is a type that benefits from elaboration but that selection information is a type that doesn't. Then writers need never worry about finding good elaborations for selection information; they can always leave selection information unelaborated no matter what the tasks are like. There might also be other types of information that do not benefit from elaboration. So, if there is reason to believe that the third explanation is correct, then our research strategy should be to explain why syntax information differs from selection information in this respect and attempt to predict what other types of information will benefit or not from elaboration.

The research reported in the following chapters attempts to sort out these alternatives by creating conditions under which people are more or less likely to need selection information and providing different types of selection elaborations. The experiment is described in detail in Chapter 4. First, however, Chapter 3 analyzes what is involved in selecting a procedure and reviews related research.

Chapter 3
THE ROLE OF SELECTION INFORMATION IN SKILL LEARNING

This dissertation focuses on the last of the three subcomponents of skill learning presented in Chapter 2: how people learn to choose appropriate procedures for solving a problem. This chapter begins by considering the problem of procedure selection and what kinds of tasks require strategies for selecting a procedure. I will then consider what information in an instructional text might be relevant for learning such strategies and briefly review what work has been done in this area. The second part of this chapter deals with what instruction on selection should look like, focusing in particular on the role of examples. The chapter concludes by previewing how the experiment to be reported attempted to manipulate the features of task and quality of elaborations that were described at the end of Chapter 2.

3.1. Selection Information and Task Characteristics

The experiments of Rieter, Charney and Morgan (in press) and Carroll (1985) both concern skills in which people learn a set of loosely connected, unordered procedures that can be combined to achieve a wide variety of goals. In this respect, learning to use a computer operating system or a computer text editor is similar to learning how to cook. A good cook knows how to use some basic procedures, such as sautéeing vegetables or making a white sauce, to prepare a

wide variety of dishes. Similarly, the commands in a typical text editor can be combined to serve goals as diverse as writing computer programs, writing poetry, doing data entry, and so on. Just as someone who is skilled in cooking can combine the basic techniques to achieve new culinary delights without depending on a recipe, so a person skilled in using a computer must be able to select and combine appropriate procedures to achieve his or her own goals.

The wide range of possible goals for using a computer system (or learning some comparable skill) has a profound impact on the content of the typical instructional text. It would be difficult if not impossible to explicitly address all the goals a user might adopt. Consequently, most computer manuals describe the commands and procedures in the abstract, so that users can apply them to whatever goals they may have adopted at the moment. One common result is that new learners develop "functional fixedness": they associate a command with whatever purpose they first used it for and forget that the command may have other valuable uses. Another common result is that people finish reading a description of a command without having the slightest clue as to when they'd ever want to use it. Even when there are examples of how to use the command or procedure, at best the example reflects a guess about what the most typical use might be, and at worst the example itself is arbitrary or formal. For example, in the IBM manual for the Disk Operating System (Anonymous, 1984), the command for renaming a file is illustrated with the following example: "The command: *RENAME B:ABODE HOME* renames the file ABODE on drive B to HOME." Computer manuals are not the only instructional texts to present arbitrary or generic examples, of course. The problems that students work on in math classes or even writing classes often bear little resemblance to the real world problems that they will need these skills to solve.

Skill learners, then, may need help connecting individual procedures to higher order goals to help them remember that there is a relevant or appropriate command for a given situation. They also need to know how individual procedures relate to each other, so that when procedures are similar, they can choose the most appropriate one for the given circumstances. I will use the term *selection information* for information in an instructional text that aims at satisfying either of these needs. The question is, to what extent should selection information be elaborated in the text? The experiments described in the previous chapter suggested such information can be conveyed with little or no elaboration. One goal of the dissertation is to test the generality of that result.

It is worth noting that selection information is probably not necessary for other kinds of skill learning, such as learning to assemble a device or operate a piece of equipment. Whereas the procedures in a computer manual can be used for a variety of goals, the goal of an assembly task or an operations task is fairly fixed. In an assembly task, such as learning to put together a stereo system, there is a specific thing that the pieces are going to form. In an operations task, such as learning to operate a radio set, there may be several different operating procedures to learn, or one general procedure with branches for various contingencies, but each procedure for the most part has a distinct goal. Because the goals are more definite, the descriptions of the procedures can also be more specific. In some cases, the steps or procedures may be presented in a fixed order that must be followed exactly. If the procedures are sufficiently complex and detailed, learners might never expect to work independently of the instructions, such as pilots who review printed check-lists each time they fly. In contrast, once learners know how to use a computer text editor, they use the manual mainly to learn new features or to

solve some unexpected problem or for an occasional reminder. For these reasons, the ability to select an appropriate procedure for the task at hand is much less important in assembly or operations tasks than in learning a system of "multi-purpose" procedures. The instructional materials for assembly and operation tasks probably don't need selection information.[7]

The work that is relevant to selection information has thus far been sparse. There is some early evidence from Smedslund (1968) that without any instruction, people can develop consistent and efficient strategies as they work through a series of problems. Not surprisingly, Smedslund found great individual differences in the quality of the strategies. Other studies suggest that you can train people to use the strategy you want them to use (e.g., Reder, in press; Sternberg & Katron, 1982). In these studies, the task instructions told the subjects which strategy to use, either for the whole experiment or for each individual problem. The situation I want to look at here is a bit different. What I'm interested in is advice that tells people "Use procedure *x* only when you are in situations that have such-and-such characteristics, otherwise use procedure *y*." The question is, when people see such advice and then are presented with a range of different situations, can they pick out the ones with the right characteristics and will they use the advised procedure?

This kind of advice is obviously very task-dependent; advice is much more limited than the general problem-solving heuristics that Polya (1957) or Newell and Simon (1972) were interested in. Advice is much more closely related to the selection rules

---

[7]Kieras (1985) and Smith & Goodman (1982) provide interesting experimental evidence that instructional texts for assembly and operation tasks do benefit from a different type of information: "how it works" information. Kieras argues that knowing how the parts of a device interact helps a learner infer procedures for how to put it together or operate it.

from Card, Moran and Newell's (1983) model of task performance. Card, Moran and Newell studied experienced users of a computer text-editor that offered two basic methods of moving the cursor: searching for a specified string of characters or moving the cursor up or down a line at a time. They found that the users had identifiable selection rules for choosing between these methods (e.g., use the search method if the target location is more than three lines away, use the line-feed method otherwise). Presumably, these computer users developed their strategies themselves, but their early learning was not observed. At least some people had developed fairly inefficient strategies. For example, one subject never used the string search method; she used some variation of the line-feed method even when the target location was over 10 lines away. . Card, Moran and Newell successfully modelled how the experts used selection rules, but they weren't interested in the relative efficiency of the rules their subjects had come up with nor in how the subjects had acquired their rules.

These issues are important for teaching novices to deal with a new set of procedures. If novices don't appreciate when to use various procedures, they might completely overlook procedures that would be very useful to them. Instead, they may settle on some inefficient procedure that they happened to learn first or that may be easy to remember. Even assuming that people do know about alternative procedures, the decision of which one to use might depend on a personal preference for solving problems a certain way, or it might depend on the specific features of the problem (i.e., perhaps people can easily guess or figure out that one method is better for a given problem). The goal of this research is to find out whether the decision-making process can be influenced by advice in the manual on when to use a specific procedure. So the work that will be described here poses two sorts of research questions:

1. Will people follow advice about when to use a command? When they don't, is it simply because the task situation itself makes it easy to identify the most efficient strategy?

2. What form should the advice take to have the most effect? In particular, should advice be stated just as a verbal rule or should it be elaborated with examples?

3.2. What It Takes to Follow Advice

As described above, the sort of advice that will be studied here is essentially a rule of the form: "Use procedure *x* only when you are in situations that have such-and-such characteristics; otherwise use procedure *y*." The advice is intended to guide a choice between two or more procedures that can be used to solve some problem; the advice points to the easiest or most efficient procedure for the circumstances.

In order to follow advice, learners must do the following things:

1. Remember the advice.
2. Be motivated to follow the advice.
3. Decide whether or not the task situation matches the conditions specified in the advice.
4. Carry out the recommended procedure.

The first requirement is straightforward; learners can't follow advice if they don't remember it. The form of the advice and the degree to which it is elaborated, may affect how well it is remembered, as will be discussed below.

The second requirement is based on the fact that, although advice is a rule, compliance with the rule is discretionary rather than compulsory. People don't have to follow advice in order to do their work; the worst that can happen if they don't is

a drop in efficiency. In contrast, syntactic rules for writing correct computer commands are compulsory in the sense that if a person wishes to issue a command, he or she must follow the rule exactly. In real-world situations, learners may be strongly motivated to work efficiently, perhaps in order to meet a deadline or conserve resources. In the present research, the nature of the task provides a different kind of motivation: the task is to find the most efficient way to solve a problem. This task forces subject to consider various solutions to a problem and compare how many steps each one would take. These comparisons will be difficult if there are many possible solutions, each with a fairly large number of steps. In that case, subjects might be more prone to trust the advice to point to the most efficient solution, rather than trying to perform all the necessary calculations.

The third requirement addresses the fact that subjects must analyze the situation in a problem to decide whether the advice applies. In fact, the advice recommends two different actions depending on the characteristics of the situation. Procedure *x* is to be used if there is a positive match between the conditions in the advice and the task situation. If there is no match, then following the advice means using procedure *y*. Recall, for example, the selection rules for moving the cursor from Card, Moran and Newell (1983): The Search method was used if the target location was more than three lines away from the current location, otherwise the Line-feed method was used. In order to use procedure *y*, learners have to recognize the absence of the conditions under which they would use procedure *x* (i.e., recognize that the cursor is *not* more than three lines away). Such "positive" and "negative" matches may affect how easy it is to decide which procedure is the advised one.

Finally, once it is clear which procedure is advised for the current situation, the learners must know how to carry it out, which they learn from procedural information such as syntax rules or step-by-step instructions.

3.3. Finding the Most Effective Form of Advice: the role of examples

A piece of advice is a rule: it points out something to do (or not to do) in a particular situation. Previous research has found that examples help people learn to apply rules in both the problem-solving and concept-learning domains. The present research examines the role of examples for learning to follow advice. Is advice more memorable or easier to apply correctly if it is elaborated with examples? Are people more willing to follow advice that is elaborated with examples? If examples do benefit learning, what is it about the examples that causes the effect? Is it that examples provide specific, concrete instantiations of the general terms in the rule? Or is it that examples provide a model of the task subjects have to perform?

This section will briefly review previous research on examples. It will also describe two types of examples and the different effects they may have for learning to follow advice.

3.3.1. The Benefit of Examples for Concept and Skill Learning

Numerous researchers have found that examples help people learn concepts. The basic finding is that studying a definition (in essence, a rule for category membership) along with examples of class members, greatly improves a learner's ability to correctly identify members of the class, as compared to learners who study the definition without examples.[6] Pollchik (1975) defined various psychological "defense mechanisms," with or without examples of what someone might do while exhibiting that defense mechanism. Subjects who studied definitions with examples were much better at classifying descriptions of behavior patterns. Nitsch (1977)

---

[6]In spirit, this work follows a long tradition of equating knowledge of a concept like lamp with the ability to recognize instances of real lamps in the world.

obtained similar results when subjects learned novel, unfamiliar concepts. Her subjects were also better at identifying novel category members if they had seen examples.[9]

Examples often present typical instances, but less typical instances have been found useful for learning the range and variation of a category or set. Nitsch (1977) found that subjects learned concepts better when the examples were different from each other than when they were all fairly similar. Counterexamples have also been found to help learners set boundary conditions to a rule or generalization (Tennyson, 1973; Tennyson, Woolley & Merrill, 1972). Tennyson, Woolley & Merrill took three features of examples into account (the typicality of the examples, the degree of similarity between superficial features of the examples and the similarity of examples and counterexamples) to successfully predict how the wrong mix of examples leads to overgeneralization, undergeneralization and misconception. In addition to the choice of examples, researchers have found that the order of the examples is important; subjects learn better when typical examples are presented before more exotic ones (Ellis & Anderson, 1984; Tennyson, Steve & Boutwell, 1975). A more thorough review of the literature on examples can be found in Mandl, Schnotz and Tergan (1984).

In addition to aiding concept learning, examples have also been found to facilitate problem solving. Pepper (1981) studied the effect of different computer programming textbooks on students' ability to do programming problems. He found that students who read a carefully written chapter that included numerous examples rated it more highly than comparable chapters that did not contain examples. More importantly,

---

[9]The discussion of the studies by Pollchik (1975) and Nitsch (1977) are based on Bransford's (1979) account.

these students also solved more programming problems correctly than students who read the other chapters.

Ross (1984) found that superficial similarities between the problem that subjects are currently working on and examples they previously saw in the instructional materials can influence their choice of procedure. For example, subjects learned a pair of procedures for using a computer text editor. In the instructional materials, one procedure was illustrated in a task involving a shopping list. The example for the other procedure involved a course listing. When subjects subsequently worked on editing a shopping list, they tended to use the procedure they had seen associated with a shopping list, even though either procedure would have worked equally well. This effect of "reminding" has potentially adverse consequences: in subsequent studies, Ross found that subjects tended to use the procedure they were reminded of, even if it was inappropriate for the problem at hand.

3.3.2. How Examples Aid Learning

Hobbs (1979) notes that an example involves a relationship between two statements: a rule (or generalization) and a specific instance for which the rule is true. Hobbs defines an example as a statement that asserts the same proposition as the rule, except that one or more general terms in the rule are replaced by specific terms describing class members. Drawing the identity relationship between the propositions may be an essential part of recognizing that a statement is an example. In an unpublished study (Charney, 1983), subjects read sentences and classified them as either examples or details. If the sentences were read in the context of appropriate generalizations, the examples were classified correctly 79% of the time. But read in isolation, without the context of the generalizations, the

examples were indistinguishable from the details; only 58% were correctly classified as examples.[18] Interestingly, it seems that seeing an example is not enough; the connection between the example and the rule must be explicit in order for the example to help learners apply the rule. Pirolli and Anderson (1985) studied subjects learning to write recursive functions in the programming language LISP. Seeing examples of recursive functions helped subjects write functions when the examples were presented with a rule as part of a discussion of how to write such functions. But the same examples had no effect on performance when presented as part of a discussion of how recursive functions work (i.e. tracing through the variable bindings and function calls).

Given this relationship between general rules and instances, there seem to be at least four ways in which examples might help people learn concepts or solve problems. Studying examples may:

1. Improve memory for the critical features of the rule;
2. Clarify general terms in the rule by illustrating the range and variation of class members;
3. Convince learners of the utility or truth of the rule;
4. Provide a basis for analogy to new problems.

Examples may improve memory for the critical features in two ways. First, as suggested above, comprehending the example as an example may involve rehearsing the relationships between the critical elements of the rule. The specific terms of the example are matched onto the general terms of the rule. To see how this works, consider the following generalization and two elaborations (taken from Charney, 1983):

---

[18]The difference in performance due to the reading context was significant for both the percent of examples that were classified correctly, $t(83) = 8.4$, $p < .01$, and for d', the subject's ability to detect the examples against the "noise" of the details, $t(43) = 8.0$, $p < .01$.

1. "Lawsuits are now pending which seek to hold handgun manufacturers and distributors liable for the damage caused by their products."

2. "The family of James Riordan, a Chicago police officer killed by a handgun, is suing Walther, the West German maker of the gun and International Armament Corporation, its American distributor."

3. "The cases are based on an unconventional and as yet unproven application of the product liability law, the law made famous by the suits against the Corvair and Pinto automobiles."

To recognize that sentence (2) is an example of the generalization in (1), but that sentence (3) is not, readers might have to realize that (1) and (3) state the same proposition, by matching the Riordan family to the initiators of a lawsuit (i.e., the agents of a suing action), matching Walther to the manufacturers and the IAC to the distributors (the objects of the suing action) and matching the death of the officer to the damage caused by the handgun (the reason for the suit). Recognizing examples may therefore provide rehearsal of the critical relationships between the arguments in the rule as well as encoding variability, both of which might improve recall. Furthermore, the extra specificity, concreteness and familiarity of the terms in the example may aid recall, since concrete and specific terms are generally recalled better than abstract terms (e.g. Paivio, Yuille & Madigan, 1968).

Even if learners remember the rule, they may not understand it well enough to apply it correctly. So examples may be an effective way to clarify and illustrate the terms of the rule. The research of Tennyson, Woolley and Merrill (1972) cited above is relevant here; without the right selection of examples, learners may overgeneralize, undergeneralize or misconceive the scope of the rule. Again, in order for examples to fulfill this function, learners must draw the necessary relationships between the terms of the example and those of the rule. Tennyson, Steve and Boutwell (1975) added analyses of the examples to a text teaching subjects to recognize metrical

forms in poetry. Each example was discussed to show how it met the critical attributes of the metrical category. The authors found that subjects who read the text with these elaborations performed better on a classification task than subjects who saw the unannotated examples.

In contrast to the preceding analysis, the work of Reder and Anderson (1980, 1982), described in Chapter 2, suggests that elaborations will impair recall of the generalization and comprehension scores. In fact, Mandl, Shnotz and Tergan (1984) essentially replicated Reder and Anderson's results. They prepared two versions of an expository text on "Man and His Environment," which differed only in that one version contained examples of the general concepts. Like Reder and Anderson, they found that recall and comprehension scores were at least as good (and sometimes better) when the text contained fewer elaborations (in this case, examples).[11]

As in Reder and Anderson's studies, performance in the Mandl et al. experiment was measured with declarative tests in which subjects had to recall or make simple inferences about the main concepts, but not apply them to solve problems. We can bring these results into line with the classification and problem-solving research cited above (where examples did improve performance) if we assume that in both classification and problem-solving tasks, examples helped people apply the rules. In the case of classification, the learner may use a rule to test whether a putative

[11]Similarly, Charney (1983) found no difference in subjects' ability to remember the generalizations, whether the generalizations had been elaborated with examples or details or studied in isolation, without elaboration.

member of a class meets the necessary and sufficient conditions for membership.[12] In the case of problem solving, such as learning to write correct computer commands, learners use the rules to generate computer code that meets constraints specified in the rules. So the value of examples for helping learners remember or understand rules (at least to the extent of being able to answer comprehension questions) is still an open question.

It also remains to be seen whether examples increase learners' aptitude for following advice. The final two features of examples listed above may be relevant to this question. First, examples may have an important role for establishing the truth or the utility of a rule (Perelman & Olbrechts-Tyteca, 1969; Schoenfeld, 1979; Mandl, Schnotz & Tergan, 1984; Gilson & Abelson, 1968). Seeing a variety of examples may convince learners that a rule is truly general. Seeing relatively easy examples may convince learners that a rule is easy to apply. This affective function of examples may be particularly important when the rule is a heuristic or a piece of advice that learners are not obliged to follow. Seeing cases where using a heuristic greatly facilitates problem solving may convince learners to use it themselves.

Finally, research in skill acquisition suggests that people who are learning math or computer programming rely heavily on examples of correctly solved problems as models for solving new problems. Pirolli and Anderson (1985) observed subjects learning the programming language LISP. Thinking-aloud protocols revealed that the subjects drew detailed analogies between the problem they were working on and

---

[12] Or, as many argue, learners decide on a putative member by drawing analogies between it and known members and nonmembers (e.g., Hintzman, 1980), or a generalized prototype (e.g., Rosch, 1977). In this case, the example is serving as a model rather than as an auxiliary aid for remembering or understanding the terms of the rule. See the discussion of examples as models below.

worked-out examples they had seen in the instructional text. The example served as a model or template for the solution. Reliance on examples as models has also been documented in other problem domains, such as learning to solve analogies (Lefevre and Dixon, 1984).

3.3.3. Conceptual Examples and Task Examples

It isn't clear which of the four aspects of examples described above are most important for skill learning. The present work takes steps toward distinguishing the instantiation function of examples from the "model" function. For this purpose, I will characterize examples that instantiate general concepts as concept examples and worked-out problems, such as those found in a math or programming textbook, as task examples. These two types of examples are similar to what Mandl, Schnotz and Tergan (1984) call "illustrative examples" and "application examples," respectively. Both types of examples provide specific instantiations of the general terms of a rule, but the types differ in what other kinds of information they provide.

To illustrate the two kinds of examples, consider the following three sentences. The first is a stylistic rule, based on Williams (1981). The second is a concept example of the rule, and the third is a task example:

1. Rule. When a nominalization follows an empty verb, change the nominalization to a verb that replaces the empty verb.

2. Concept Example. For example, nominalizations such as *investigation, inquiry* or *response* often follow empty verbs such as *make* or *conduct*. Use the verbs *investigate, inquire* or *respond* instead.

3. Task Example. For example, change the sentence "The police conducted an extremely thorough investigation into the incident," to "The police investigated the incident extremely thoroughly."

The concept example instantiates the general terms "empty verb" and

"nominalization," but doesn't provide a context in which they might occur. The task example instantiates the general terms within a specific context. The context illustrates something about the situations in which the rule should apply: the nominalization need not follow the empty verb directly. It also illustrates something about how to carry through the solution: changing a noun to a verb can necessitate changes to other parts of the sentence.

If instantiation is the major contribution of an example, then concept examples and task examples should aid performance to the same degree. But if it is important to use the example as a model, then seeing a task example should improve performance more than seeing a concept example. Task examples may also help people remember a rule when they are working on a task, because seeing the task may remind them of the example (Ross, 1984). Finally, the task example may be better for demonstrating the utility of the advice, by showing rather than just asserting that following the rule leads to a desirable outcome.

3.4. A Preview of the Experiment

The purpose of the experiment is to discover whether or not elaborating on selection information improves learners' performance, and if so, what sort of elaboration is most effective. At the end of Chapter 2, three features were described that may control whether elaborations have a chance to improve performance: the tasks must be ones for which the information is relevant, the elaborations must be the right kind and the information must be a type that benefits from elaboration. Of these three features, only the nature of the tasks and the nature of the elaborations can be varied experimentally. If we can find a task in which some type of selection elaborations improve performance, then obviously selection information is a type that can benefit from elaboration.

The experiment to be reported here creates conditions under which selection information should be highly relevant to good task performance. A set of problems was constructed which could be solved with various combinations of procedures. Subjects were instructed to find the most efficient combination of procedures to solve each one. Some subjects received advice on when a particular procedure was most efficient and some did not. If such selection information is indeed relevant to the tasks, then subjects who see advice should perform better than those who see no advice.

In an attempt to increase the likelihood that subjects follow the advice, some problems were designed to be more difficult than the others. The rationale for this manipulation is as follows. Since advice in our definition is discretionary, subjects may decide not to follow it; they may try to find the most efficient solution some other way, such as counting the number of steps in each possible solution and comparing them until they have found the most efficient one. If so, then we may find no effect of advice or elaborations simply because subjects prefer to use their own method or because either method allows subjects to find the right answer equally easily. To anticipate this possibility, half of the problems were designed to be more difficult; these problems have a greater number of possible solutions, each with a greater number of steps. Subjects may then be unable to mentally compute a solution by counting and comparing steps. So while subjects might be content to compute solutions for the easy problems, they may have to fall back on the advice for the hard ones.

If this analysis is correct, then the evidence that advice benefits from elaborations may appear only for the hard tasks. That is, subjects might not follow the advice on easy tasks, so there would be no difference in performance for subjects who have

seen some form of advice and those who haven't. However, on hard tasks, subjects should be much more likely to benefit from the advice. And if elaborations are needed in order for subjects to take full advantage of the advice, then subjects who see advice elaborated with examples should perform much better than either subjects who see unelaborated advice or subjects who see no advice. If seeing examples makes no difference to performance even on hard tasks, then we may have found another situation in which selection information fails to benefit from elaborations.

The nature of the elaborations will be manipulated in this experiment by providing either concept examples or task examples to illustrate the advice. For the reasons described in the previous section, examples have been found to be very effective aids for learning to apply rules. Since advice is a type of rule, we would expect examples to be the best kind of elaboration. However, it is not certain what kind of example would be most effective. The section on examples above presented some arguments for why task examples might be more effective than concept examples.

An additional factor of interest is how tenaciously subjects adhere to advice. Ideally, we want students to follow advice judiciously, to use it as a recommendation rather than as a commandment. In this experiment, we will be able to gauge how often and how eagerly subjects follow the advice in two ways. First, we will see whether subjects follow the advice on easy tasks (when they might be able to compute the solution independently), as well as hard tasks. The experiment also employs a more direct measure of how often subjects follow the advice. This measure relies on the fact that advice is a heuristic rule rather than an algorithmic rule. This means that advice is not guaranteed to lead to the desired outcome. To reflect this possibility, some problems were designed for which following the advice would lead to an incorrect solution to the problem. If subjects follow the advice

closely, they will answer incorrectly on these problems. That is the advice even when it leads them astray.

subjects may follow

Chapter 4
METHOD

4.1. Overview of the Box-World Game

The experiment involved a game called "Box-World," in which simple geometric objects were displayed on a Dandyllon computer. Each play of the game presented a *situation* (a configuration of boxes and objects on the screen) and a *goal* for what objects particular boxes should contain (see Figure 4-1).[13] The goal could be achieved with a combination of commands for moving objects from one box to another or changing an object's shape. In one task, the Efficiency task, the object of the game was to figure out the most efficient combination of moves and changes for achieving the goal. Subjects performed this work mentally; they didn't actually issue the commands to the computer. When subjects believed they had arrived at the most efficient solution, they signalled the computer by pressing a key. The computer then proposed an action (see Figure 4-2) for achieving part of the goal and subjects had to decide whether or not this action was part of the most efficient way to achieve the goal. They signalled their decision by pressing either a key labelled "yes" or one labelled "no."

The most efficient solution path to a goal is determined by the total number of

---

[13]Because the reproductions of the Dandyllon screen image were of poor quality, the Figures depicting the Box-World game were prepared on an Apple Macintosh.

FIGURE 4-1

BOX-WORLD

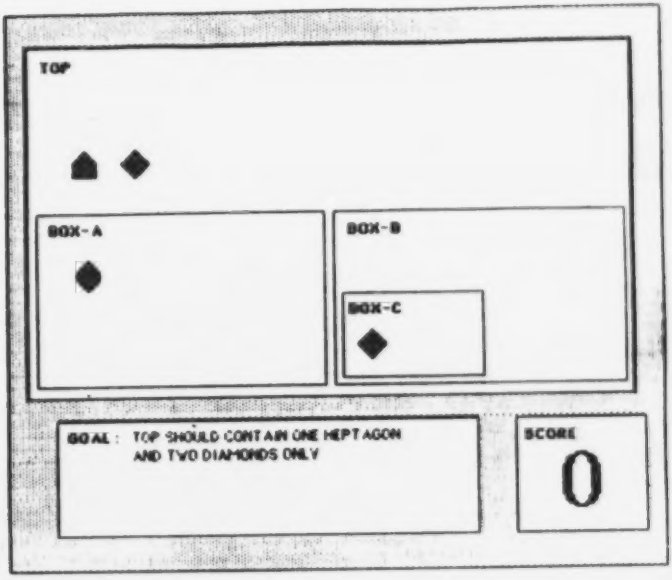
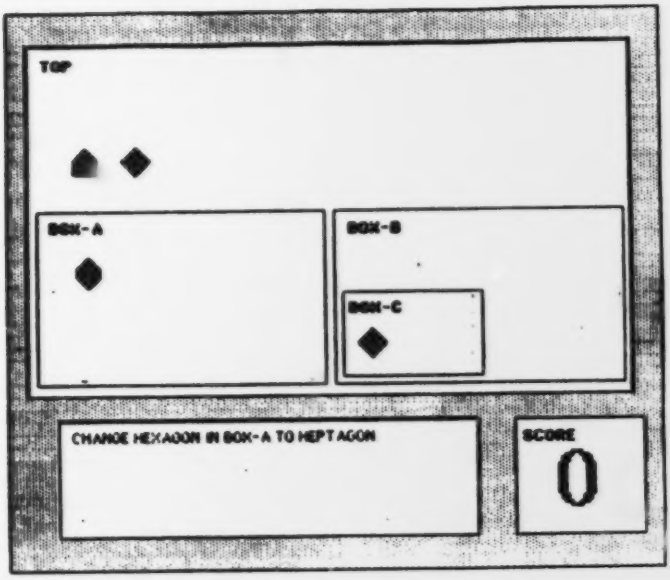


FIGURE 4-2

BOX-WORLD



individual Change and Move commands required. In order to find the most efficient solution path, subjects had to take into account some restrictions on shape changes and movements. One set of restrictions affected how many commands would be required to change an object to another shape. In particular, the five kinds of Box-World objects (triangle, diamond, pentagon, hexagon and heptagon) were ranked according to how many sides they have (i.e., three, four, five, six and seven sides, respectively). A single CHANGE command can change an object only one degree up or down in rank. So, for example, a diamond can be changed to either a triangle or a pentagon with a single command. However, a sequence of two commands would be required to change it to a hexagon, or three commands to change it to a heptagon. Similarly, the movement of an object from a given box was restricted to a box nested directly above or below the current box. So a sequence of MOVE commands would be required to move an object to a more distant box. These restrictions are explained more fully in the Box-World manual, provided in Appendix A.

4.2. Design

The experiment included two tasks using the Box-World game that employed somewhat different designs, an Efficiency task and an Advice Recognition task. The Efficiency task required subjects to find the most efficient solution to a problem. The Advice Recognition task required them to decide whether or not a proposed action was consistent with the advice.

4.2.1. Efficiency Task

The Efficiency task required subjects to find the most efficient solution to a problem. This task employed a 4x2x2 mixed factorial design. The first factor, Instruction, was a between-subjects factor that manipulated the availability and form of advice in the instructional materials that subjects studied. Four versions of the materials were available: *No Advice*, *Rule Alone*, *Rule plus Concept Example* and *Rule plus Task Example*. Subjects were randomly assigned to study one of these versions, with the constraint that the four groups contain an equal number of subjects.

The second factor, Appropriateness, was a within-subjects factor that manipulated the types of trials in the Box-World game. Three-quarters of the trials were *Appropriate* and one-quarter were *Inappropriate*. On Appropriate trials, following a strategy consistent with the advice in the manual led to the most efficient solution (and so the correct answer). On Inappropriate trials, this strategy led to an incorrect response. The purpose of this factor was to measure how apt subjects were to follow the advice; the closer they followed the advice, the more likely it was that they would respond correctly on Appropriate trials and incorrectly on Inappropriate trials.

The third factor, Difficulty, was also within-subjects. For both levels of appropriateness, half the trials were designed to be *Easy* and half were *Difficult*. Difficulty was determined by two features: the number of plausible solution paths for achieving the goal and the number of steps in the most efficient solution path.

The parameter of interest was the subject's decision about whether an action proposed by the computer was part of the most efficient solution path. So the dependent measures were the accuracy of these decisions and how long subjects took to respond.

4.2.2. Advice Recognition Task

The Advice Recognition task required subjects to decide whether or not a proposed action was consistent with the advice (regardless of whether it was part of the most efficient solution). This task employed a 3x4x2 mixed factorial design. The first factor was again the between-groups factor Instruction. Instruction had three levels for this task, namely the three versions of the manual that contained some form of advice (*Rule Alone*, *Rule plus Concept Example* and *Rule plus Task Example*). Subjects were assigned to an Instruction condition before completing the Efficiency task and kept the same assignment for the Advice recognition task.[14]

The second factor, Advice-Response Match, was a within-subjects factor of trial characteristics. There were four trial categories: *Use-Yes*, *Use-No*, *Don't Use-Yes*, *Don't Use-No*. Trials were categorized as Use or Don't Use according to whether the advised strategy would dictate using a particular command for that situation or avoiding that command. Trials were categorized as Yes or No according to whether the proposed solution was consistent with the advice or not. Overall, there were the same number of Use trials as Don't Use, and the same number of Yes's as No's. This factor is described more fully in the Materials section, and examples of the trials are included in Appendix B.

The third factor was Difficulty; as in the Efficiency task, half the trials were *Easy* and half were *Difficult*.

The dependent measures were accuracy of response and reaction times.

---

[14]Subjects in the *No Advice* group also performed the Advice Recognition task, in order that all subjects would be treated equivalently. These subjects received special instructions for the task that presented a form of the advice. However, the data from the No Advice group were not included in the analyses for this task.

4.3. Materials

4.3.1. Instructional Materials for Box-World

The instructional materials consisted of a 4-page manual for the Box-World program (reproduced in Appendix A). The manual briefly introduced the Box-World domain and then described the Move, Change and Delete commands. This basic manual represented the No Advice instructional condition. Three other versions of the manual were prepared that differed only by the addition of some form of advice for how to use the Change or Move commands efficiently. One manual simply stated the advice as a general rule; the other two manuals added different kinds of examples to illustrate the advice. The three forms of advice were:

- Rule Alone: Advice stated as a general rule without any elaboration.
- Rule plus Concept Example: Advice stated as a rule with a verbal example giving specific instantiations of the terms in the rule.
- Rule plus Task Example: Advice stated as a rule with an annotated pictorial example of a task situation showing that following the advice leads to the most efficient solution.

Figure 4-3 shows these three forms of advice for using the Change command. To create the manuals for the three advice conditions, a sheet containing the appropriate form of advice was inserted after the description of the Change command in the basic manual.

The Move Isomorph. The advice told subjects under what conditions to use a particular procedure and when to avoid using it. Since the manual described two major procedures (the Change command and the Move command), it was important to vary which procedure the advice concerned, while holding constant as many of the task features as possible. Accordingly, a second set of advice was prepared that concerned the Move command, shown in Figure 4-4.

FIGURE 4.3

Three forms of Advice for the Change Command

a. RULE ALONE

CHANGE is usually the most efficient command to use whenever you can make an Object into the shape you want by issuing just one command; otherwise avoid using CHANGE.

---

b. RULE PLUS CONCEPT EXAMPLE

CHANGE is usually the most efficient command to use whenever you can make an Object into the shape you want by issuing just one command; otherwise avoid using CHANGE.

EXAMPLE: Use CHANGE when you have a diamond and need a triangle or a pentagon; you can get to either of these shapes with just one command. Look for another way to solve the problem if you have a diamond but need a hexagon or a heptagon.

C. RULE PLUS TASK EXAMPLE

CHANGE is usually the most efficient command to use whenever you can make an Object into the shape you want by issuing just one command; otherwise avoid using CHANGE.

EXAMPLE: Consider the following Box-World situation. Suppose you want BOX-A to contain two triangles. Changing DIAMOND2 into a triangle takes only one command, and is more efficient than changing PENTAGON2 into a triangle (2 changes) or moving TRIANGLE3 in BOX-B to BOX-A (2 moves).

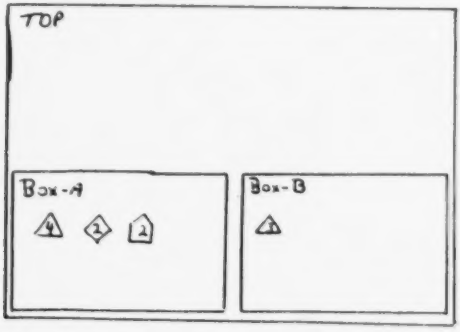


FIGURE 4-4

Three forms of Advice for the Move Command

a. **RULE ALONE**

MOVE is usually the most efficient command to use whenever you can put the Object into the Box you want by issuing just one command; otherwise avoid using MOVE.

---

b. **RULE PLUS CONCEPT EXAMPLE**

MOVE is usually the most efficient command to use whenever you can put the Object into the Box you want by issuing just one command; otherwise avoid using MOVE.

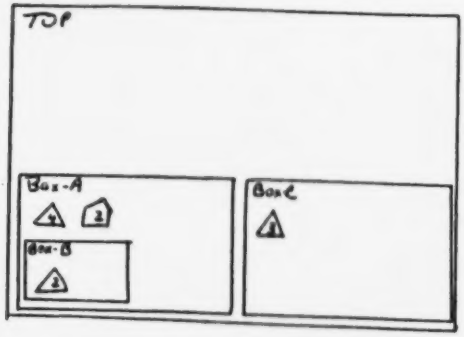
EXAMPLE: Use MOVE when you have a diamond in BOX-A and need a diamond in either BOX-B (which is inside BOX-A) or in TOP (which contains BOX-A); you can get to either of these boxes with just one command. Look for another way to solve the problem if you need a diamond in any other Box.

---

c. **RULE PLUS TASK EXAMPLE**

MOVE is usually the most efficient command to use whenever you can put the Object into the Box you want by issuing just one command; otherwise avoid using MOVE.

EXAMPLE: Consider the following Box-World situation. Suppose you want BOX-A to contain two triangles. Moving TRIANGLE2 from BOX-B to BOX-A takes only one command, and is more efficient than moving TRIANGLE3 from BOX-C to BOX-A (2 moves) or changing PENTAGON2 into a triangle (2 changes).



In order to balance the location of the advice, the order of topics in the manuals was controlled; when the manual contained advice about the Change command, the discussion of the Change command followed that of the Move command. When the advice concerned the Move command, the discussion of the Move command followed that of the Change command. Separate versions of the No Advice version of the manual, representing each topic order, were also prepared.

4.3.2. Trials for the Box-World Game

A pool of 64 Box-World problems was prepared, each consisting of a situation, a goal and a proposed action that was coded for a correct response (either Yes or No). The construction of the trials involved four factors: Appropriateness, Difficulty, Use and Response (see below). The four factors were completely crossed, yielding a 4x2x2x2 design (the first factor represents the distribution of three Appropriate trials for every Inappropriate trial). Consequently, 32 trials were needed for one complete replication and the 64-item pool represented two complete replications.

This pool of items was used for both the Efficiency task and the Advice Recognition task. A fixed set of 32 problems was used in the Advice Recognition task. The problems were chosen randomly from the pool of 64 problems, with the constraint that the set of problems for each task should represent one complete replication of the experimental design. The only modification needed for using the problems in the Advice Recognition task was an adjustment in the coding of the correct response.

Each Box-World Situation presented four to ten objects arranged inside any of two to six boxes. The Goal specified for one or more of the boxes what objects (if any) they should contain. The Action statement described an action performed on

just one of the objects, either changing its shape or moving it to another box. (See Figures 4-1 and 4-2.) The specified change or movement in the Action was not always consistent with the advice. That is, the Action statement might propose changing a diamond to a heptagon or moving an object to a distant box. The Action was always part of a plausible solution to the problem and was unique to one solution path.

Appropriateness. Plausible solutions were computed for each Situation-Goal combination. If there was no one solution that had fewer steps than the others, then the problem was discarded. If the most efficient solution required violating the advice (e.g., it called for changing a diamond to a heptagon), then the problem was classified as Inappropriate. If the best solution did not violate the advice, then the problem was characterized as Appropriate. Three-fourths of the trials in the final pool of problems were Appropriate and one-fourth were Inappropriate.

Difficulty. Difficulty was determined by two factors: the number of plausible solution paths for achieving the goal and the number of steps in the most efficient solution path. There was an average of 3.7 plausible paths per problem for easy problems and 4.4 paths for hard problems. The most efficient path for easy problems averaged 3.8 steps; for hard problems, the average was 4.4 steps.

Use/Don't Use. The Use factor varied whether the advised strategy would dictate using a particular command to achieve a goal or looking for an alternative to using that command. The advice recommended using the Change command under certain circumstances, namely whenever an object could be changed into the desired shape by issuing just one command. This is the positive form of the advice, telling the subject to use the change command when the situation met certain conditions. The

advice also has a negative aspect: it says to avoid using the Change command in all other circumstances, but it doesn't say what to do instead. Half of the trials in the pool satisfied the conditions in the advice; that is, there was an object that could be changed with one command to a shape specified in the goal. For the other trials, there was no object that met these conditions.

Response. The Response factor insured that there were an equal number of trials in which the correct response was *Yes* and the correct response was *No*.

The trials in the Advice Recognition task were classified into four categories according to the Use/Don't Use dimension and the Yes/No dimension: *Use-Yes*, *Use-No*, *Don't Use-Yes*, *Don't Use-No*. The most straightforward application of the advice is expected to come in either the Use-Yes or the Don't Use-No categories. In the Use-Yes trials, the advice dictates using the Change command and the action statement proposes the relevant change, so the correct response is Yes. In the Don't Use-No trials, the advice recommends against using the change command (for instance, when there are no objects available that can be changed with just one command into the desired shape). The action statement proposes a change that violates the advice, so the correct response is No.

Samples of trials in the different conditions are found in Appendix B. Note that subjects saw the trials in the format used in Figures 4-1 and 4-2. For ease of reference, the samples in the appendix list the goal and action and are annotated with the correct responses for both the Efficiency task and the Advice Recognition task and explanations of how the trial satisfies the requirements of the conditions.

4.3.3. The Move Isomorph

It was important to hold as many task features constant as possible when the advice concerned the Move command. However, following the Move advice to solve a problem would not lead to the same solution as following the Change advice. Therefore, isomorphs of each of the 84 problems were created (these will be referred to as Move *isomorphs*). The Move isomorphs were created by systematically transforming the situation, goal and action of each problem so that the solution paths and the number of steps per path would be preserved and so that following the Move advice would lead to a solution that was formally identical to that in the corresponding Change isomorph.

Creating the Move isomorphs involved a translation between *distance of movement* and *degree of shape change*. Every step in the solution path of a problem that involved a Change command was translated into a step involving a Move command, and vice versa. For example, suppose that one method of solving a problem would involve changing a diamond into a hexagon. Given the constraints on the Change command described above, this method would involve two steps (or two Change commands). In the Move isomorph of this problem, changing the shape twice translates into moving the diamond to the desired location from two boxes away. The method still involves two steps, this time two Move commands. The move isomorphs of the examples from Appendix B are found in Appendix C.

Because this translation preserved all of the relevant characteristics of the problems, the set of 84 Move isomorphs fulfilled all the requirements of the experimental design. In addition, the sampling of problems chosen for the Efficiency task and Advice Recognition tasks was preserved. That is, the subjects who studied

the Move advice worked on the Move isomorphs of the 32 problems chosen for the Advice Recognition task, while subjects who studied the Change advice worked on the Change isomorphs of these problems. Differences due to whether subjects studied advice about the Move command or the Change command will be referred to henceforth as the isomorph factor.

4.4. Apparatus

The experiment was conducted on Xerox 1108 Dandylion or Dandytiger computers with 17-inch, bit-mapped, high resolution displays (1024 x 808 pixels). Software was developed to record and timestamp the subjects' responses.

4.5. Subjects

The subjects were 113 students and staff members at Carnegie-Mellon University and the University of Pittsburgh. All subjects were native speakers of English, or fluent enough to completely understand the manual. Subjects received a basic compensation of either money ($3.50) or class participation credit. In addition, all subjects were paid a bonus of five cents for each correct response they made above chance; the highest possible bonus was $1.60.

4.6. Procedure

One to five subjects were run concurrently at individual Dandylion workstations. In the first phase of the experiment, subjects were randomly assigned to either the Change or Move isomorph and to one of the four instruction conditions. They were given their assigned version of the manual and a fixed period of five minutes in which to read it. Subjects were told to review the information if they finished reading before the five minutes were up, since the manual would not be available to

them while they worked on the tasks. After the study phase, subjects were asked to do the Efficiency task, a Recall task and the Advice Recognition task. Subjects performed the tasks in this order so that performance on one task would not contaminate the results of subsequent tasks.

An initial group of 48 subjects (composed primarily of CMU students) did not do the Advice Recognition task; they performed the Efficiency task on all 64 problems. A second group of 65 subjects (primarily Pitt students) did 32 trials with the Efficiency task instructions and 32 trials with the Advice Recognition instructions. Differences between these two groups of subjects will be referred to henceforth as the Replication factor; the first group will be referred to as the Eff-Only group and the second as the Eff&Adv group.

4.8.1. Efficiency Task

The exact instructions for this task are presented in Figure 4-5.

For each trial, subjects were presented with a situation and a goal (Figure 4-1). Subjects were instructed to study the situation for as long as they liked to figure out the most efficient combination of moves and shape changes to achieve the goal. When they were ready, they pressed the space bar. The goal statement disappeared and was replaced by the action statement (Figure 4-2). Subjects decided whether or not the action was part of the most efficient solution. They signalled their response by pressing either a key labelled "yes" or a key labelled "no." The computer then gave a feedback message telling the subject whether the response was correct or not. If the response was correct, the score in the lower righthand corner of the screen was increased by one. There were five practice trials, during which subjects were allowed to ask questions about the procedure.

FIGURE 4-5

Instructions for the Efficiency Task

Now you will play a game using Box-World. In each play of the game, you will see some Boxes and Objects on the screen and a goal for what objects you want certain boxes to contain. The computer will propose an action that may or may not be part of the most efficient way to get to the goal. (The proposed action will never be enough to get to the goal all by itself, but it may be one of the things you would want to do). The object of the game is to decide whether the proposed action is part of the most efficient strategy for reaching the goal, that is the strategy using the smallest combination of the Box-World commands you read about. If you think the proposed action is part of the most efficient strategy, press the key labelled "Yes." If you think it's not, press "No." Each time you are correct, you score a point. Each time you are wrong, the computer scores a point. As your Bonus, we will give you a nickel for every point you score, but subtract a nickel from the Bonus for every point the computer scores.

There are two important things to know about the actions that the computer will propose. (1) This action will describe either the final position or the final shape of one of the objects. Carrying out this action may take one or more separate commands. The most efficient solution is based on the smallest number of separate commands it would take to reach the goal. (2) The proposed action may be to change the shape of an Object. To reach the goal, you might have to move this Object too. Similarly, you might have to change the shape of an object that the computer proposes to move. As long as you think this sequence is part of the most efficient solution to the problem, say "Yes." If it's not, say "No."

Here is the procedure for each play:

(1) Press the space bar. The computer will display a situation and a goal (printed on the bottom part of the screen).

(2) Press the space bar again when you are ready to see the proposed action. The goal statement will disappear and will be replaced by the proposed action.

(3) Keep your index fingers resting on the YES and NO keys. Press the key labelled YES if you think that the proposed action is part of the most efficient way to get to the goal. Press the key labelled NO if that action is not part of the most efficient solution.

(4) The computer will tell you whether or not your decision was right and update the score.

(5) To start the next play, press the space bar.

We want you to play as quickly as possible and still get a high score. In order to let you get used to the procedure, there will be five practice plays first. You will be able to take a break half-way through the game.

For each trial, the computer recorded the response, whether it was correct or incorrect, and two response time intervals:

- *Encoding and planning interval*: the time that elapsed between the initial presentation of the situation and when the subject pressed the space bar.
- *Decision interval*: the time between the presentation of the action statement and when the subject signalled a decision.

The computer generated a different random presentation order for each subject.

4.6.2. Recall Task

Subjects were asked to write short answers to 1-3 questions. The exact questions are presented in Figure 4-6. The number of questions that subjects answered depended on which version of the manual they had studied. Subjects who had seen examples in their manuals (i.e., the Rule plus Concept Example and the Rule plus Task Example conditions) answered Question 1, which asked them to recall the example. Subjects who had seen any form of advice (i.e., the two example conditions and the Rule Alone condition) answered Question 2, which asked them to recall the advice. All subjects answered Question 3, which asked subjects to retrospect on the strategy they had used to solve the problems.

4.6.3. Advice Recognition Task

The exact instructions for this task are presented in Figure 4-7.

In this task, subjects were first reminded of the advice they had seen in the manual. They were shown the page of the manual with the advice they had studied, including examples in the appropriate conditions. Then the subjects performed a series of 32 recognition trials. In each trial, subjects were presented with a situation and a goal (Figure 4-1). Subjects studied the situation for as long as they liked.

When they pressed the space bar, the goal disappeared and was replaced by the proposed action (Figure 4-2). Subjects decided whether or not the action was consistent with the advice and signalled their response by pressing either a key labelled "yes" or a key labelled "no." The computer then gave a feedback message telling the subject whether the response was correct or not. If the response was correct, the score in the lower righthand corner of the screen was increased by one. There were five practice trials, during which subjects were allowed to ask questions about the procedure.

As for the Efficiency Task, the computer recorded the response for each trial, whether it was correct or incorrect, and the two response time intervals. The computer generated a random order for presenting the trials for each subject.

FIGURE 4-6

Questions in the Cued-Recall Task

1. The manual for Box-World that you read at the beginning of the experiment offered some advice and an example about when to use the commands. Write down what you remember of the example.

2. The Box-World manual offered some advice about when to use the commands. Write down what you remember of the advice.

3. Describe the strategy you used to solve the problems.

FIGURE 4-7

Instructions for the Advice Recognition Task

This final part of the experiment involves another game. In this part of the experiment, we are not interested in whether the proposed action is part of the most efficient way to get to the goal. Instead, we want you to decide whether or not the proposed action follows the advice you read in the manual about when to use the commands. In order to remind yourself of the advice, you may now review that page of the manual (attached).

The procedure for playing the game is similar to the previous game, except that instead of trying to find the most efficient solution, you are simply deciding whether the proposed action follows the advice or not. In other words, does the proposed action accomplish part of the goal in a way that is consistent with the advice? Press the key labelled "Yes" if you think the proposed action is consistent with the advice, even if it is not part of the most efficient solution. Press "No" if you think the action is not consistent with the advice. Each time you are correct, you score a point. Each time you are wrong, the computer scores a point. As your Bonus, we will give you a nickel for every point you score, but subtract a nickel from the Bonus for every point the computer scores.

Here is the procedure for each play:

(1) Press the space bar. The computer will display a situation and a goal (printed on the bottom part of the screen).

(2) Press the space bar again when you are ready to see the proposed action. The goal statement will disappear and will be replaced by the proposed action.

(3) Keep your index fingers resting on the YES and NO keys. Press the key labelled YES if you think that the proposed action is consistent with the advice. Press the key labelled NO if that action is not consistent with the advice.

(4) The computer will tell you whether or not your decision was right and update the score.

(5) To start the next play, press the space bar.

Again, we want you to play as quickly as possible and still get a high score.

Chapter 5
RESULTS AND DISCUSSION, PART I: THE EFFICIENCY TASK

5.1. Overview

The central task in this experiment is the Efficiency task. The results from this task speak to the following four key questions:

1. Did subjects follow the advice when searching for the most efficient solution to a problem?
2. Was the advice helpful? That is, were subjects who followed the advice able to identify the most efficient solution more consistently or more quickly than subjects who didn't see any advice?
3. Did the difficulty of the task affect the subject's reliance on the advice? Further, did the difficulty of the task interact with the need for elaboration? That is, were elaborations of the advice more effective when the tasks were more difficult?
4. What form of advice was most effective? Did the advice need to be elaborated with examples, and if so, what sort of example was most effective?

These questions are of primary importance since they address the ways in which advice and elaborations might influence a person's selection strategies during actual problem solving. The other two tasks, the Recall and Advice Recognition tasks, attempt to delve more deeply into what the subjects learned from reading the different forms of advice. Recalling the advice and recognizing actions that are consistent with it seem to be reasonable prerequisites to following the advice. So

these tasks were intended to check that subjects learned enough to meet these prerequisites and to reveal any differences between the forms of advice.

The results and discussion will be presented in two chapters. This chapter focuses on the Efficiency task and treats each of the four key questions above in turn. The next chapter presents the results from the Recall task and the Advice Recognition task, as well as some results that limit the generality of the findings.

5.1.1. A Note on the Number of Subjects per Task

The data reflect the scores of a total of 113 subjects, but these subjects did not all perform all three tasks:

- As described above, 48 subjects (the Eff-Only group) performed 64 trials of the Efficiency task without doing the Advice Recognition task at all.
- For 18 subjects in the Eff&Adv group, only the Efficiency task data is available; although they performed both tasks, their Advice Task data were thrown out after a programming error was discovered (they received incorrect feedback on approximately 20% of their trials). The Efficiency Task data for these subjects were retained, and 17 additional subjects were run on both tasks as replacements.
- While 12 subjects in the No Advice group did perform the Advice Recognition task (see the Procedure section), their data was not included in the analysis.
- Finally, three subjects inadvertently failed to complete the recall test.

Appendix D shows the total number of subjects who completed each task as a function of instructional group and problem isomorph (Change or Move). The parenthesized entry for each task is the number of subjects in the Eff-Only replication group.

5.2. The Utilization and Utility of the Advice: the Efficiency Task

The Efficiency task assessed the extent to which subjects followed a strategy consistent with the advice in order to find the most efficient solution to a problem. Subjects were asked to decide whether a proposed action was part of the most efficient solution. Their performance was assessed in terms of the accuracy and speed of these decisions. For each of the Efficiency task measures, a 4x2x2x2 analysis of variance (ANOVA) was performed over the Instruction, Appropriateness, Difficulty and Replication factors. The Change and Move isomorphs were analyzed separately and will be discussed separately because the cell means pointed to quite different patterns of behavior. The following discussion pertains only to results from the Change isomorph.

Table 5-1 shows what percentage of the subjects' decisions were correct as a function of the Instruction, Appropriateness and Difficulty variables. The data are presented separately for the Easy trials (the top of the table) and Hard trials (the bottom). The table indicates that subjects were sensitive overall to differences in the types of trials. Subjects were significantly more accurate on Appropriate trials than on Inappropriate trials; on the average, subjects' decisions were correct on 78% of the Appropriate trials but only on 67% of the Inappropriate trials, $F(1,40) = 17.4$, $p < .01$. Subjects were also significantly more accurate on Easy trials (77% correct), than on Hard trials (68% correct), $F(1,40) = 17.4$, $p < .01$.

There was no overall effect of Instruction, suggesting that none of the instructional manuals led subjects to perform much better or much worse overall than any of the others. The lack of an effect of Instruction is suprising, since it implies that no form of advice influenced behavior significantly compared to the control group, the subjects

TABLE 5-1

EFFICIENCY TASK:

Percent Correct Decisions as a Function of Instruction, Appropriateness and Difficulty Change Isomorph

	INSTRUCTION				
	No Advice	Rule Alone	Concept Example	Task Example	
EASY TRIALS					MARG.
Appropriate	.87	.83	.80	.80	.83
Inappropriate	.74	.63	.59	.86	.71
MARG.	.81	.73	.70	.83	
HARD TRIALS					
Appropriate	.73	.79	.73	.68	.73
Inappropriate	.65	.58	.58	.70	.63
MARG.	.69	.69	.66	.69	

who saw no advice. The more detailed analyses below will attempt to sort out just where advice had an effect and where it didn't.

5.3. Did Subjects Follow the Advice?

The signal for whether or not subjects followed the advice was the Appropriateness effect: following the advice should have misled subjects into making incorrect decisions on the Inappropriate trials. The effect of Appropriateness reported above suggests that all subjects performed better when the trials were Appropriate.[18] However, a significant interaction between the Instruction and Appropriateness factors ($F(3,40) = 4.5$, $p < .01$) suggests that this effect is mainly due to two instructional groups, the Rule Alone group and Concept Example group. Scores for subjects in the Rule Alone group averaged about 20 points higher when trials were Appropriate than when they were Inappropriate, for both Easy and Hard trials ($t(11) = 2.8$, $p < .05$ and $t(11) = 3.9$, $p < .01$, respectively). Similarly, scores for the Concept Example group averaged 20 points higher on Easy Appropriate tasks than on Easy Inappropriate tasks ($t(11) = 2.8$, $p < .05$) and about 15 points higher on Hard Appropriate tasks than on Hard Inappropriate tasks, although the latter contrast did not reach significance.

The difference between Appropriate and Inappropriate trials for the other two groups of subjects was smaller or non-existent. The No-Advice group was less accurate on the Inappropriate trials, but the differences were only on the order of 10 points, and the contrasts were not significant for either the Easy trials or the Hard

---

[18] Note that the main effect of Appropriateness is not due to any artifact of trial construction or selection. The mean number of solution paths and steps per path in the Inappropriate trials were the same as the means for the Appropriate trials.

trials.[18] The drop-off in accuracy on the Inappropriate trials may simply mean that some subjects in the No Advice group devised strategies that were similar to the advice. Surprisingly, the Task Example group showed no signs of following the advice: their scores were even slightly *more* accurate on the Inappropriate trials. (An explanation for this result will be offered below).

5.4. Was the Advice Helpful?

Even though there is evidence that at least some subjects followed the advice, it turns out that in the present experiment, the advice was not very helpful. As the accuracy scores in Table 5-1 suggest, there was no overall advantage to seeing the advice. Subjects in the No Advice group were just as accurate (or more so) than the subjects who saw advice. The lack of an overall advantage for the advice would not be important if it could be attributed exclusively to the Inappropriate trials (where subjects were deliberately misled into making mistakes). However, it is apparent from the data for the Appropriate trials that subjects who saw advice did not identify the most efficient solution more often than subjects without advice, even when that solution was consistent with the advice.

It is not surprising that the No Advice group did so well on the Easy trials, since these trials were designed to be simple enough that subjects could compute the solutions on their own. What is surprising is that the advice failed to improve performance on the Hard trials, for which performance was low overall. Nevertheless, subjects who saw advice were no more accurate on the Hard Appropriate trials than the subjects who saw no advice.

---

[18] Other results also support the claim that effect of Appropriateness is due to the Rule Alone and Concept Example groups and not the No-Advice group. A separate ANOVA was performed on partitioned data, omitting the data for the No Advice group. The main effect of Appropriateness and the Appropriateness x Manual Interaction were both still significant ($F(1,30) = 12.2$, $p < .01$ and $F(1,30) = 8.9$, $p < .01$, respectively.)

The lack of a gain in accuracy might have been offset by a gain in speed of response. However, the reaction time measure revealed no overall speed advantage associated with following the advice. Table 5-2 shows the average amount of time (in seconds) that subjects spent on the Encoding and Planning interval, studying the situation and goal of a trial. The data are presented as a function of Instruction, Appropriateness and Difficulty.[17]

Overall, subjects spent about 38 seconds per trial. As in the percent correct measure, there was no main effect of Instruction, but subjects' times were sensitive to the type of trial. There was a main effect of Difficulty; subjects were 14 seconds faster on Easy trials than Hard trials, $F(1,40) = 181.4$, $p < .01$. There was also a smaller main effect of Appropriateness; subjects were 8 seconds faster on Appropriate trials than Inappropriate trials, $F(1,40) = 22.2$, $p < .01$.

The lack of a main effect of Instruction again suggests that there was no overall advantage to studying the advice. The Concept Example group was in fact the fastest group, taking about 10 seconds less per trial than the No Advice group on both Easy and Hard trials. However, these differences were not statistically significant. Furthermore, the Rule Alone group, which was the other group that apparently followed the advice, turned in relatively slow times, especially on the Hard trials.

No advantage of advice emerged either for the second reaction time measure, the Decision interval. Table 5-3 shows the average amount of time (in seconds) that

---

[17]This and all other reaction time measures to be reported reflect times for correct responses only. In order to moderate the effect of extremely fast and extremely slow times, the reported ANOVAs were performed on the log transforms of the subjects' times.

TABLE 5-2

EFFICIENCY TASK:

Encoding and Planning Time (in seconds), as a function of Instruction, Appropriateness and Difficulty Change Isomorph

	INSTRUCTION				
	No Advice	Rule Alone	Concept Example	Task Example	
EASY TRIALS					MARG.
Appropriate	30.0	29.5	25.0	26.6	27.8
Inappropriate	38.4	34.6	23.8	31.8	32.2
MARG.	34.2	32.1	24.6	29.2	
HARD TRIALS					
Appropriate	43.8	46.3	35.4	42.1	41.9
Inappropriate	47.5	60.6	37.7	47.9	48.4
MARG.	45.7	53.5	36.5	45.0	

subjects took to make a correct decision. The data are presented as a function of Instruction, Appropriateness and Difficulty. For this measure, there were no main effects of Instruction, Appropriateness or Difficulty.[18] Overall, subjects took about 7 seconds to make a decision. Reaction times in the Rule Alone and Concept Example conditions were each about 1.3 seconds faster than the No Advice group, but the differences again failed to reach statistical significance.

In sum, we must conclude from the evidence available that the advice was not necessary to good performance on this task, since subjects in the No Advice group were just as accurate and just as quick to respond to Appropriate trials as subjects who studied the advice. Obviously there is no value to having advice that is only right part of the time, that doesn't improve accuracy (compared to having no advice) when it is right and that significantly lowers accuracy when it is wrong. Why didn't the advice work better? It is not that subjects could easily figure out the solutions on their own; the poor performance on the Hard trials suggests that subjects needed help of some sort. So the problem seems to lie either with the advice itself or with the subjects' ability to follow it. The former possibility suggests that the advice itself needs to be changed, the latter that we haven't yet found the right way to train people to follow the advice.

One possible problem with the advice itself is that it didn't always lead uniquely to the most efficient path. That is, there was often more than one solution path for a trial that was consistent with the advice. So while studying the problem and trying

---

[18]There was a small three-way interaction of Instruction x Appropriateness x Difficulty, $F(1,40) = 2.8$, $p = .05$. This interaction seems to be due to the No Advice and Rule Alone groups which were faster on Appropriate trials when the trials were Easy, but faster on Inappropriate trials when the trials were Hard.

TABLE 5-3

EFFICIENCY TASK:

Decision Time (in seconds), as a function of Instruction, Appropriateness and Difficulty Change Isomorph

	INSTRUCTION				
	No Advice	Rule Alone	Concept Example	Task Example	
EASY TRIALS					MARG.
Appropriate	7.5	5.5	6.3	7.0	6.6
Inappropriate	8.3	7.9	6.9	7.0	7.5
MARG.	7.9	6.7	6.6	7.0	
HARD TRIALS					
Appropriate	8.5	6.8	6.5	6.9	7.2
Inappropriate	7.8	6.2	6.8	7.5	7.1
MARG.	8.1[a]	6.5	6.7	7.2	

to find the most efficient path, subjects may have occasionally discovered a path that was consistent with the advice, unaware that it was not the most efficient path, and looked no further. When presented with an action, they responded on the assumption that the path they had chosen was the right one. When the trials were Appropriate, these subjects would be able to correctly reject actions that proposed violating the advice, but they might make two kinds of mistakes. They might incorrectly accept actions from the inefficient path they had chosen or incorrectly reject actions from the path that was actually most efficient. On the other hand, all these subjects would be consistently misled on the Inappropriate trials because following the advice would still lead them to ignore solutions that blatantly violated the "one change" restriction, even though these solutions turned out to be most efficient. Although this interpretation is consistent with the pattern of results, it would be difficult to verify it without knowing what solution path subjects believed was the most efficient. If this interpretation is correct, however, then the advice is no good; it fails to narrow the search space of solutions sufficiently to help subjects find the best solution.

There is also the possibility that subjects did not follow the advice closely enough. This possibility will be discussed in section 5.6 below and in the next chapter.

Even if the advice is determined to be bad advice, the results of this experiment are interesting for what they tell us about how people use advice (good or bad): how heavily they rely on it and whether different forms of advice or different tasks affect the degree of this reliance.

5.5. Did Task Difficulty Affect Reliance on Advice?

As reported above, subjects were significantly more accurate on Easy trials than on Hard trials, and needed significantly less time in the Encoding and Planning interval for Easy trials than Hard trials. However, Difficulty did not interact with Appropriateness on any measure: making tasks harder did not seem to increase subjects' reliance on the advice. As the accuracy results in Table 5-1 indicated, subjects in the Concept Example and Rule Alone groups showed evidence of following the advice even on the Easy trials, when they should have been able to compute the solutions themselves. Furthermore, these subjects were not more apt to follow the advice on the Hard trials; the spread between the scores for Appropriate and Inappropriate trials was again about 18 points.

One might be tempted to conclude that these subjects followed the advice blindly, never attempting to analyze the problems on their own. However, there is indirect evidence that these subjects retained a certain degree of independence. In particular, following the advice blindly would have led subjects to miss *all* of the Inappropriate trials, but performance on the Inappropriate trials never dropped below chance, even on the Hard trials.

Perhaps the reason that subjects followed the advice to the same degree for both Easy and Hard trials is that these trials were mixed together. That is, trials were presented in random order and subjects may have faced a Hard task at any point in the experiment. So if subjects received some Hard trials early in the experiment, they may have decided to use the advice to solve them. Once they had thought to use the advice, though, they may have stuck to the same method even for Easy trials. This interpretation can be tested by counterbalancing the order in which subjects see trials of different levels of difficulty.

5.6. Did the Form of the Advice Affect Performance?

This section focuses on the differences between the groups that studied the advice. First, I will attempt to explain the puzzling behavior of the Task Example group, which unlike the Rule Alone and Concept Example groups showed no signs of following the advice. In particular, I will argue that the Task Example group may have interpreted the example in an unexpected way that led them to adopt a strategy that was different from the advice. Second, I will compare the performance of the Concept Example and Rule Alone groups to see what effect the example may have had on performance.

5.6.1. The Effect of the Task Example

Subjects in the Task Example condition showed no sign of following the advice while performing the Efficiency task. The data presented in Table 5-1 suggest that these subjects were even slightly more accurate on the Inappropriate trials than the Appropriate trials. Two explanations of this behavior seem especially plausible:

1. Subjects didn't devote the necessary time to working through the Task Example. They skimmed over both the advice and the example.[19]
2. Subjects adopted a new strategy based on the example, which inadvertently focused attention away from the advice itself.

According to the first explanation, subjects didn't follow the advice because they couldn't remember it. The Task Example was the lengthiest form of advice. Given the five-minute time constraint on reading the manual, and not knowing that efficiency would be important for the tasks, subjects may have felt this information was

---

[19] Alternatively, subjects may have had difficulty processing the verbal example in conjunction with the pictorial diagram, since these tasks call on visual processes to handle different kinds of information simultaneously. Presenting the example orally while subjects view the diagram may solve this problem.

relatively safe to ignore. Therefore, while performing the Efficiency task these subjects never called on whatever weak mental representation of the advice they might have formed. To anticipate the data to be reported in Table 5-1 below, subjects in the Task Example group actually recalled the advice much worse than subjects in the other groups.

According to the second explanation, subjects didn't remember the advice because they interpreted the example in an unexpected way. Consequently, they either forgot the advice or reinterpreted it to conform to the example. Then, during performance of the Efficiency task, they followed a strategy based on their interpretation of the example. LeFevre and Dixon (1984) and LeFevre (1985) found exactly this sort of behavior in their research on instructions for solving analogy problems. They found that when verbal instructions for how to solve a problem (i.e., rules) are contradicted by a task example, people tend to follow their interpretation of the example. How might subjects have reinterpreted the Task Example in the present experiment? Consider again the Task Example in Figure 4-3. It compares three different solutions to the problem and shows that the one consistent with the advice requires the fewest steps. The most likely new interpretation of the example is that it's always necessary to carry out a systematic path-length comparison, like the one in the example. This interpretation is quite different from the advice. The advice was intended as a short-cut so that subjects would not have to compute and compare the steps of all possible paths. Instead subjects should only have looked for paths that used the Change or Move commands in a particular way.

The reinterpretation explanation is supported to a certain extent by the dissimilarity of the percent correct data for the Task Example group and the No Advice group in Table 5-1. If the Task Example group had simply ignored the advice

altogether, the results of these two groups should have been quite similar. In fact, the two groups were equally accurate overall, but the Task Example group tended to be more accurate on the inappropriate trials. This result is consistent with the proposed reinterpretation: If the Task Example subjects routinely computed all path lengths, then they shouldn't have been misled on the inappropriate trials. There is also some evidence that is inconsistent with the reinterpretation explanation. If the Task Example group was the only group to compute multiple path lengths, then one might expect their reaction times in the Encoding and Planning interval (Table 5-2) to be longer than any other group. In fact, their times were not the longest.[20]

Further research will be necessary to decide between these two explanations. One way to decide the question would be to require a criterion level of recall on all parts of the manual before allowing subjects to perform the Efficiency task. If good recall of the advice is the key, then Task Example subjects should show more evidence of following the advice. An alternative method would be to take thinking-aloud protocols of subjects reading the manual and performing the Efficiency task. The protocols should reveal whether the Task Example subjects ignore the Advice or whether they interpret it differently than the other subjects.[21] If the reinterpretation explanation proves correct, then it will be interesting to investigate whether subjects can be trained to follow the advice with modified task examples or perhaps a combination of task and concept examples.

---

[20]The results from the Advice Recognition task, to be presented in section 6.2, are consistent with either interpretation. Before performing this task, subjects were explicitly reminded of the advice and the example. The results suggest that if the Task Example subjects had remembered the advice, or had interpreted it correctly, they would have been able to apply it just as well as the other groups.

[21]Unfortunately, the subjects' retrospective reports on their solution strategies (see section 6.1.3 below) were too vague to shed any light on this question.

5.6.2. The Effect of the Concept Example

The Rule Alone and Concept Example groups both showed evidence of following the advice for the Efficiency task. As shown in Table 5-1, the two groups were about equally accurate: overall the Rule Alone group was correct on 71 percent of the trials, and the Concept Example group on 65 percent. The Concept Example group appeared to work more quickly than the Rule Alone group (see Table 5-2), taking an average of 12 seconds less per trial. However, the difference in reaction times was only marginally significant ($p < .10$) for Hard trials.

There was a difference, though, in sensitivity to the Appropriateness factor, as reflected in how quickly subjects worked. The Concept Example group seemed oblivious to the Appropriateness manipulation: the average difference between the overall reaction times for the Appropriate and Inappropriate trials was only 1 second. In contrast, the Rule Alone group spent an average of 10.5 seconds longer on the Inappropriate trials than on the Appropriate trials. Most of this extra time came in the Encoding & Planning interval for the Hard trials. The contrast performed on the differences between the times for the Appropriate and Inappropriate trials was significant, $t(12) = 2.7$, $p < .05$.

So, the Concept Example subjects spent less time on the Encoding & Planning interval, especially on Hard Inappropriate trials. At this point, it is only possible to speculate about why this difference occurred. One possibility is that the Concept Example minimized the number of alternative solution paths that subjects considered while looking for the most efficient path. In particular, the Concept Example people may never have considered paths that clearly violated the advice, even though these were in fact more efficient for the Inappropriate trials. The Rule Alone group may

have begun to consider other paths, especially when these paths looked reasonably efficient, as in the inappropriate trials. The extra time the Rule Alone spent did not lead them to abandon the advice; so they must have ended up either trusting the advice more than their own calculations or giving up on the alternative solutions too soon.

Why might the Concept Example have disinclined subjects from doing the same thing? The Concept Example gave concrete, specific instances of both "legal" and "illegal" changes (i.e., "Use the Change command if you have a diamond and need a triangle or a pentagon. Avoid using Change if you have a diamond and need a hexagon or heptagon.") Seeing the negative instances may have made it easier for subjects to identify violations of the advice (Tennyson, 1973). When subjects saw the problem situation and goal, they might have automatically set a limit of what shapes to consider changing and ruled out of consideration any objects that would necessitate violating the advice. One way to find out more about the effect of the Concept Example would be to take thinking-aloud protocols and see whether the Rule Alone group considers more alternative solutions or a different set of alternatives from the Concept Example group.

It would be interesting if the Concept Example influenced which solution paths subjects were willing to consider. All in all, however, the effect for the example is fairly minor. Subjects apparently did not need the example in order to follow the advice, and having the example did not lead to significantly faster reaction times. This result (or non-result) is surprising since examples have been shown repeatedly to aid rule application. Taken as a whole, the results suggest that Concept Examples are not an effective form of elaboration for advice. The strongest conclusion that could be drawn from the similarity between the Concept Example and Rule Alone

groups is that we have found a second, more compelling instance of selection information, in the form of advice, failing to benefit from elaborations. However, until the effects of the Task Example are sorted out, it is too soon to draw such a conclusion. It may be that in order to take advantage of the advice, subjects need to use a worked-out solution (such as a Task Example) as a model. However, as suggested above, the Task Example in this experiment may have been inadequately processed or misinterpreted.

5.7. The Effect of the Move Isomorph

The results presented above concerned subjects in the Change isomorph only. It is now time to consider the differences between the Change and Move isomorphs. The intention of creating the Move isomorph was to vary the topic of the advice while keeping as many of the task features constant as possible. So the Move isomorph subjects read advice about the Move command and Change isomorph subjects read advice about the Change command. Ideally, the topic of the advice should have been irrelevant to the effects of the advice. However, the results for the Efficiency task were somewhat different for the Move isomorph.

Table 5-4 shows the percentage of correct decisions as a function of Instruction, Appropriateness and Difficulty. As in the Change isomorph, subjects were sensitive to the difficulty of the trials. Subjects' scores on the Easy trials were about 9 points higher than on the Hard trials, $F(1,57) = 31.8$, $p < .01$. Overall, accuracy on Appropriate trials was slightly higher than on Inappropriate trials, especially for Hard trials. However, Appropriateness did not produce a significant main effect or interact with any other variable.

The results for the No Advice and Task Example groups are very similar to those

TABLE 5-4

EFFICIENCY TASK:

Percent Correct Decisions as a Function of Instruction, Appropriateness and Difficulty Move Isomorph

	INSTRUCTION				
	No Advice	Rule Alone	Concept Example	Task Example	
EASY TRIALS					MARG.
Appropriate	.82	.71	.66	.87	.77
Inappropriate	.75	.71	.72	.85	.76
MARG.	.79	.71	.69	.86	
HARD TRIALS					
Appropriate	.76	.64	.64	.76	.70
Inappropriate	.62	.61	.58	.69	.63
MARG.	.69	.63	.61	.73	

in the Change isomorph. However, the accuracy of the Concept Example and Rule Alone groups is suddenly much lower than the other Instruction groups. Reflecting the relatively poor performance of these two groups, the Instruction factor produced a significant main effect, $F(3,57) = 3.4$, $p < .05$. The Concept Example and Rule Alone groups in the Move isomorph differed in another way from their counterparts in the Change isomorph. Whereas in the Change isomorph these groups showed the biggest effect of Appropriateness, in the Move isomorph the Concept Example and Rule Alone groups showed little or no drop in accuracy on the Inappropriate trials.

The reaction time measures provide little additional information of interest. Table 5-5 shows the average amount of time (in seconds) that subjects spent in the Encoding and Planning interval and Table 5-6 shows the average amount of time (in seconds) in the Decision interval. The data in both tables are presented as a function of Instruction, Appropriateness and Difficulty. For both measures, the only factor that produced a significant effect was Difficulty. In the Encoding and Planning interval, subjects averaged about 3.2 seconds on Easy trials as opposed to 4.4 seconds on Hard trials, $F(1,57) = 62.8$, $p < .01$. In the Decision interval, subjects spent about 5.8 seconds on Easy trials and 7.2 seconds on Hard trials, $F(1,57) = 19.8$, $p < .01$. Thus, the measures themselves were again sensitive enough to pick up the expected differences among types of trials. In addition, while the times for the No Advice and Task Example groups are comparable to the times for the corresponding groups in the Change isomorph, the Concept Example group is no longer fastest on either measure.

While it is disturbing that the two isomorphs did not yield identical effects of advice, the fact of a difference in performance for the two isomorphs is not in itself too remarkable. Hayes and Simon (1977) found that a subject's ability to solve the

TABLE 5-5

EFFICIENCY TASK:

Encoding and Planning Time (in seconds), as a function of Instruction, Appropriateness and Difficulty
Move Isomorph

	INSTRUCTION				
	No Advice	Rule Alone	Concept Example	Task Example	
EASY TRIALS					MARG.
Appropriate	31.0	25.7	31.5	36.3	31.1
Inappropriate	31.5	29.8	33.7	34.4	32.4
MARG.	31.3	27.8	32.6	35.4	
HARD TRIALS					
Appropriate	40.9	36.6	42.1	47.2	41.7
Inappropriate	49.8	39.0	44.6	48.8	45.6
MARG.	45.4	37.8	43.4	48.0	

TABLE 5-6

EFFICIENCY TASK:

Decision Time (in seconds), as a function of Instruction, Appropriateness and Difficulty
Move Isomorph

	INSTRUCTION				
	No Advice	Rule Alone	Concept Example	Task Example	
EASY TRIALS					MARG.
Appropriate	5.0	4.9	6.4	7.3	5.9
Inappropriate	5.5	4.3	7.2	5.6	5.7
MARG.	5.3	4.6	6.8	6.5	
HARD TRIALS					
Appropriate	5.7	5.8	7.7	8.2	6.9
Inappropriate	6.0	7.3	8.0	8.3	7.4
MARG.	5.9	6.6	7.9	8.3	

Tower of Hanoi problem varies enormously depending on the isomorph. They have found the same variation for many other kinds of problems. So the important question is what caused the difference between the Change and Move isomorphs. In this regard, it is quite interesting that the same pair of instructional groups stood out in both isomorphs. The major difference between the results for the two isomorphs was the performance of the Rule Alone and Concept Example groups. In the Change isomorph, these groups showed the biggest effect of Appropriateness, but their overall accuracy was not significantly different from the Task Example and No Advice groups. In the Move isomorph, the Rule Alone and Concept Example groups showed virtually no effect of Appropriateness. Furthermore, they were much less accurate than the No Advice and Task Example groups. It seems reasonable to attribute this conjunction of results to the same cause: that these groups of subjects followed or attempted to follow the advice. However, when subjects in the Move isomorph attempted to follow the advice, they unexpectedly encountered severe problems.

Under this interpretation, the difference between the two isomorphs was caused by some difference in the advice or how subjects carried it out. Conceivably, the difference might have been caused by some flaw in the construction of the Move isomorph problems (see the description of trial construction in Section 4.3.3). It is important therefore to note that the difference is confined mainly to the Concept Example and Rule Alone groups in the Efficiency task. As described above, performance in the Task Example and No Advice groups was comparable in the two isomorphs. As we will see in the next chapter, subjects in the two isomorphs recalled the advice equally well. We will also see that, in the Advice Recognition task, overall accuracy was quite similar in the two isomorphs.[22] The fact that the

[22]Reaction times did differ for the two isomorphs. In particular, the Concept Example group was the fastest group in the Change isomorph, but the slowest group in the Move isomorph.

difference is isolated to the Rule Alone and Concept Example groups in the Efficiency task suggests that the difference between the isomorphs was caused by differences in the advice as it related to the Efficiency task and not by the transformation of the problems themselves.

So what aspects of the advice might have caused the difference? Consider again the advice for the two isomorphs, reproduced below with their respective Concept Examples.

* CHANGE is usually the most efficient command to use whenever you can make an Object into the shape you want by issuing just one command; otherwise avoid using CHANGE.

  EXAMPLE: Use CHANGE when you have a diamond and need a triangle or a pentagon; you can get to either of these shapes with just one command. Look for another way to solve the problem if you have a diamond, but need a hexagon or a heptagon.

* MOVE is usually the most efficient command to use whenever you can put the Object into the Box you want by issuing just one command; otherwise avoid using MOVE.

  EXAMPLE: Use MOVE when you have a diamond in BOX-A and need a diamond in either BOX-B (which is inside BOX-A) or in TOP (which contains BOX-A); you can get to either of these boxes with just one command. Look for another way to solve the problem if you need a diamond in any other Box.

The Change advice (and especially the Concept Example) is quite concrete and deals with simple relationships between familiar geometric objects. There's a fair degree of certainty about which shapes have the relationship specified in the advice. A diamond is always "one side more" than a triangle and "one side less" than a pentagon. In contrast, the Move advice depends on a relationship of location, involving boxes in varying configurations. The example involves one hypothetical configuration of boxes with particular names, but there is no certainty in the Box-World domain that Box-A will always contain a smaller Box-B. So the concept of an

object being "one box away" is less concrete and less definite than the concept of an object having "one more side or one fewer side." As a result, the Rule Alone and Concept Example subjects in the Move isomorph may have formed a very imprecise representation of the advice. The fact that their accuracy was so much lower than the other groups suggests that they actually relied on some representation of the advice. The fact they were equally accurate on Appropriate and Inappropriate trials suggests that whatever representation they did form was quite different from what was intended.

It should be possible to test the hypothesis that the difference in representations of the advice was due to concreteness. A more "concrete" Move isomorph could be created. Locations might be specified as slots in certain fixed positions rather than as relative positions in a nesting of boxes.[23] If the concreteness is the key, then performance in this new isomorph should be much more like performance in the Change isomorph. (Alternatively, it might be possible to create a more abstract Change isomorph in which objects take on arbitrary attributes instead of familiar geometric shapes. Then the performance in the Change isomorph should become more like the present Move isomorph).

---

[23]Thanks to Dick Hayes for this suggestion.

Chapter 6
RESULTS AND DISCUSSION, PART II

The Recall and Advice Recognition tasks attempt to delve more deeply into what the subjects learned from reading the different forms of advice. Recalling the advice and recognizing actions that are consistent with it seem to be reasonable prerequisites to following the advice. So these tasks were intended to check that subjects met these prerequisites and to reveal any differences between the forms of advice.

6.1. Recall Task

After completing the Efficiency Task, subjects were asked to write short answers to three questions (the exact questions appear in Figure 4-6).

6.1.1. Recall of Advice

Subjects who had seen any form of advice (Rule Alone, Concept Example or Task Example) were asked to recall the advice itself.

Scoring

Two judges independently scored the recall answers on the following three-point scale:

- No credit (0 points) for no answer or for answers that mentioned nothing that was relevant to how to use the commands efficiently.

- Partial credit (.5 points) for answers that mentioned efficient use of commands, capturing the spirit of the advice but omitting or incorrectly stating the "at most one command" constraint.
- Full credit (1 point) for answers that correctly stated the "at most one command" constraint.

Sample responses for each of these three coding categories are provided in Appendix E. The correlation between the two judges' scores was $r = .77$, and disagreements were resolved to mutual satisfaction.

Table 6-1 presents the mean recall scores as a function of Instruction and Isomorph. A 3x2x2 Analysis of Variance (ANOVA) was performed over the Instruction, Isomorph and Replication factors. Recall overall was not outstanding, with the mean score about .46.

The form of instruction in the manuals had a significant effect on how well subjects remembered the advice, $F(2,71) = 4.3$, $p < .05$. Subjects in the Concept Example and Rule Alone conditions remembered the advice about equally well, but subjects in the Task Example group had much worse recall (.27). As mentioned earlier, the poor recall of the Task Example group supports either of two interpretations: subjects in this group may never have processed the advice well enough to recall it or the example may have led them to reinterpret the advice in such a way that they could not recall the original interpretation accurately. Subjects in the Concept Example group, at least in the Change isomorph, appear to have recalled the advice best, but the contrast between this group and the Rule Alone group was not significant.

There was no main effect of Isomorph: subjects in the Change and Move isomorphs recalled the advice equally well overall. There was also no interaction of

instruction with Isomorph, even though subjects in the Change Isomorph seemed to have better recall than subjects in the Move Isomorph for the Concept Example condition (.71 vs. .50) and worse recall in the Task Example condition (.19 vs. .35).

Given that overall recall appeared fairly weak on this measure, it is worth reconsidering whether any subjects were able to follow the advice on the Efficiency task. The evidence that the Concept Example and Rule Alone groups did follow the advice is the fact that these groups showed the predicted effect of Appropriateness, at least in the Change Isomorph. The only difference between the Appropriate and Inappropriate trials was that on the Inappropriate trials, it was necessary to violate the advice to find the most efficient solution. Furthermore, the recall results are consistent with the Efficiency task results: the groups that showed most evidence of following the advice also obtained the highest recall scores, and the group that showed no evidence of following the advice had worst recall. This conjunction of results makes it seem reasonable to conclude that subjects in the Concept Example and Rule Alone groups were able to follow the advice, despite their relatively poor showing on the recall task.

It is also interesting to speculate on whether the poor showing of the advice in general was due to poor recall of it overall. If recall of the advice had been greater, would subjects in the advice conditions have performed better than subjects in the No Advice group? In order to indirectly assess the importance of recall on performance in the Efficiency task, the accuracy data for the three advice groups in the Change Isomorph were reanalyzed using the recall scores as covariates. However, there was no significant effect of recall on the accuracy scores, and adjusting the scores to reflect recall did not produce noticeable differences in the cell means. So as a group, the subjects in, for example, the Rule Alone group

produced a fairly similar pattern of results on the Efficiency task regardless of whether individual subjects in the group later received a high score or a low score on the recall test. What this suggests is that the recall test was not a sensitive enough measure of what subjects got out of reading the advice; some subjects who remembered and acted on the advice may have scored poorly on the recall test, perhaps because the question (or probe) was too vague.

6.1.2. Recall of Examples

Subjects who had seen examples in conjunction with the advice (i.e., the Concept Example and Task Example conditions) were asked to recall the example. Recall of the example was overall quite poor. Only 18% of the subjects recalled all or part of the example they had seen. An additional 22% of the subjects reproduced the general rule for the advice but not the specific example. Nearly a third (27%) of the subjects were unable to give any answer at all, and another 35% gave answers that didn't relate to the advice at all: they reproduced the syntactic rules for the Change and Move commands or examples of these rules.

Subjects were asked to answer this question before trying to recall the advice. Without this context, they may have gotten confused about whether the example in question was the examples for the syntax rules that were also included in the manual or the example for the advice. In addition, it may have been difficult to frame a response about the Task Example, since it involved a specific diagram and goal situation.

TABLE 6-1

RECALL OF ADVICE:

Mean Recall Scores as a function of
Instruction and Isomorph

	INSTRUCTION			
	Rule Alone	Concept Example	Task Example	
				MARG.
Change Isomorph	.53	.71	.19	.48
Move Isomorph	.48	.50	.35	.44
MARGINALS	.51	.61	.27	

6.1.3. Retrospective Report on Solution Strategy

Finally, subjects were asked to retrospectively describe the strategy they used to solve the problems. There was quite a bit of variation in the type of response subjects gave to this question. A number of responses essentially recapped the task instructions. About 60 responses focused on finding a complete solution strategy, about half of these mentioned counting steps to find the most efficient solution, the other half just referred to comparing "my solution" to "the computer's solution." Another group of about 30 subjects seemed to focus on the individual boxes rather than on an overall solution. They mentioned trying to "solve one box at a time." Of these, about half explicitly mentioned strategies for choosing between the Change and Move commands. For example, some mentioned looking only at nearby objects, or trying to move objects before considering changing shapes, or making minimal shape changes. Only eight subjects claimed to have no strategy apart from guessing or guessing on trials they thought were hard. The responses in general were somewhat vague and were not analyzed beyond this rough categorization.

6.2. Advice Recognition Task

This task assessed subjects' ability to judge whether or not a solution was consistent with the advice in the manual. Before performing this task, subjects were explicitly reminded of the advice. As in the Efficiency task, subjects were presented with a situation and a goal. After they studied the situation and the goal, the computer proposed an action for achieving part of the goal. Subjects were told to decide whether or not this action was consistent with the advice, regardless of whether the action was part of the most efficient solution to the problem. See Chapter 4 for a more complete description of this task.

Subjects' performance was assessed both in terms of accuracy and reaction times (including both how long they took to encode and plan and how long they took to make a correct decision once the Action statement appeared). For each dependent measure, a 3x4x2x2 ANOVA was performed over the Instruction, Advice-Response Match, Difficulty and Isomorph factors.

8.2.1. Accuracy of Decisions

Table 8-2 shows the percentage of correct decisions as a function of Isomorph, Instruction, Advice-Response Match, and Difficulty. The data on the left side of the table are from subjects in the Change Isomorph divided into the three Instruction conditions. The data on the right side are for the Move Isomorph. The data are presented separately for Easy trials (the top of the table) and Hard trials (the bottom).

The data show no overall effect of Instruction; subjects in all three Instruction conditions were correct on about two-thirds of the trials. There was also no effect of Isomorph; subjects who worked on the Change and Move Isomorphs were equally accurate overall. As in the Efficiency task, subjects performed significantly better on Easy trials than on Hard trials, $F(1,29) = 16.7$, $p < .01$. Subjects' scores averaged twelve percentage points higher when trials were Easy.

Although there was no overall effect of Instruction, examples did improve performance somewhat. Subjects who saw an example (either a Concept Example or a Task Example) were 10% more accurate than subjects who saw the advice without elaboration (the Rule Alone condition), $t(33) = 2.04$, $p < .05$. This result contrasts with the results of the Efficiency and Recall tasks, in which there was little difference between the Rule Alone and Concept Example groups, but both of these groups

TABLE 6-2

ADVICE RECOGNITION TASK:

Percent Correct Decisions as a Function of Isomorph, Instruction, Match and Difficulty

	CHANGE ISOMORPH			MOVE ISOMORPH		
	Rule Alone	Concept Example	Task Example	Rule Alone	Concept Example	Task Example
EASY TRIALS						
Use-Yes	.75	.71	.79	.50	.75	.83
Use-No	.54	.67	.54	.60	.71	.67
Don't Use-Yes	.58	.71	.71	.65	.67	.92
Don't Use-No	.79	.96	.79	.50	.71	.92
MARG.	.66	.76	.71	.56	.71	.84
HARD TRIALS						
Use-Yes	.50	.68	.67	.70	.50	.56
Use-No	.42	.42	.17	.40	.58	.42
Don't Use-Yes	.50	.42	.75	.50	.58	.67
Don't Use-No	.75	.83	.83	.60	.81	.75
MARG.	.54	.59	.61	.55	.62	.60

differed from the Task Example group. The example groups may have performed better on the Advice Recognition task because all subjects were reminded of the advice (and their example, if any) before completing this task and the task explicitly required comparing the proposed action to the advice. This reminder gave the Task Example group an opportunity and a motive to process (or reinterpret) the advice and the example. Since the Advice Recognition task was more of a classification task, either kind of example helped performance.

Performance on this task was relatively low, as the marginals in Table 6-2 indicate. Performance overall averaged about 67% correct. The division of trials into the Advice-Response Match categories shows that some kinds of trials were much more difficult than others, $F(3,87)=9.7$, $p<.01$. The best performance (77% correct) came on the Don't Use-No trials. In these trials, subjects were presented with an action statement that violated the advice and the correct response was No (i.e., this action was not consistent with the advice). The worst performance (52% correct) came on the Use-No trials. In these trials, there is a possible step that achieves a part of the goal and that is consistent with the advice, but this step is not the action that the computer proposes, so the correct response is No.[24]

The data concerning the Advice-Response Match factor indicate that these task features influence how easy it is to identify the advised solution. It was relatively easy for subjects to reject actions that violated the advice, but harder for them reject actions that didn't explicitly violate the advice, but didn't exploit it either. Since these task features cut across the Appropriate/Inappropriate dimension in the Efficiency task, subjects may not have been able to follow the advice well enough to

---

[24]For a sample Don't Use-No trial, see trial C-7 in Appendix B. Trial B-4 is a Use-No trial.

identify solution paths consistent with it. This may mean that subjects need more explicit instructions or "training" in how to follow the advice.

Another reason why scores in the Advice Recognition task may have been relatively low is that subjects may have confused the instructions for this task with the instructions for the Efficiency task that they had already completed. Although they were explicitly told to judge the proposed action for consistency without regard for whether it led to the most efficient solution, subjects may not have been able to drop the efficiency criterion. The effect of such "contamination" can be assessed by counterbalancing the order of the two tasks.

6.2.2. Reaction Time

The differences between the types of trials discussed above were not reflected in the reaction times. Table 6-3 shows the average amount of time (in seconds) needed to make a correct decision on an Advice Task trial (the sums of the times for the Encoding & Planning and the Decision intervals). The data are again presented as a function of Instruction, Advice-Response Match and Difficulty. The only factor to produce a main effect was Difficulty; subjects spent about 25 seconds on an Easy trial and 37 seconds on a Hard trial, $F(1,29) = 74.5$, $p < .01$.

There was, however, a significant Instruction x Isomorph interaction, $F(2,29) = 3.6$, $p < .05$, and a three-way Instruction x Isomorph x Difficulty interaction, $F(2,29) = 10.7$, $p < .01$. These interactions indicate that the Concept Example group performed differently in the Change and Move isomorphs. In the Change isomorph, the Concept Example group gave the fastest responses, while in the Move isomorph, the Concept Example group gave the slowest responses. This difference was larger for Hard tasks than for Easy tasks. This result may reflect the fact that the Concept

TABLE 6-3

ADVICE RECOGNITION TASK:

Overall Reaction Time (in seconds), presented as a function of Isomorph, Instruction, Match and Difficulty

	CHANGE ISOMORPH			MOVE ISOMORPH		
	Rule Alone	Concept Example	Task Example	Rule Alone	Concept Example	Task Example
EASY TRIALS						
Use-Yes	26.1	23.9	23.6	31.8	29.7	22.4
Use-No	28.4	26.2	26.5	20.1	38.6	22.6
Don't Use-Yes	22.6	17.2	26.3	21.1	33.3	22.0
Don't Use-No	20.0	18.3	29.6	23.2	32.5	22.8
MARG.	24.3	21.4	26.5	24.1	33.5	22.5
HARD TRIALS						
Use-Yes	39.1	28.2	42.6	32.5	48.1	39.1
Use-No	46.2	27.1	31.0	12.5	61.0	38.4
Don't Use-Yes	36.7	30.9	49.2	24.8	49.8	25.9
Don't Use-No	36.3	26.2	42.2	38.1	54.2	36.3
MARG.	39.6	28.1	41.3	27.0	53.3	34.9

Example was more concrete in the Change isomorph than in the Move isomorph (see the discussion in the previous chapter). The Task Example groups were presumably unaffected because the diagram provided extra specificity.

6.3. Replication effects

The generality of the results of this experiment is weakened by replication effects. Two distinct groups of subjects performed the experiment. The Eff-Only Replication subjects (who performed the Efficiency task but not the Advice Recognition task) were primarily students at Carnegie-Mellon. The Eff&Adv subjects (who performed both tasks), were primarily from the University of Pittsburgh. The Replication factor was included in all analyses for the Efficiency and Recall tasks. (All subjects who performed the Advice Recognition task were from the same Replication). In general, subjects in the Eff-Only Replication were quicker and more accurate than the Eff&Adv subjects. Below is a summary of the replication effects.

In the Efficiency task, subjects in the Eff-Only Replication were more accurate. Their scores were about 10 points higher than the Eff&Adv group in both the Change isomorph and the Move isomorph ($F(1,40) = 12.0$, $p < .01$ and $F(1,57) = 4.7$, $p < .05$, respectively). In the Move isomorph, there was also a significant three-way interaction of Instruction x Difficulty x Replication, $F(3,57) = 4.7$, $p < .01$. Two Instruction conditions (Concept Example and Task Example) seemed to contribute most to this interaction. Both groups were much less accurate on Hard trials in the Eff-Only Replication, but in the Eff&Adv Replication, there was no difference between Easy and Hard trials for these Instruction groups.

The reaction time measures for the Efficiency task also produced some effects of replication, but only in the Change isomorph. In the Encoding and Planning interval,

there was no main effect of Replication. There was a three-way interaction of Instruction x Difficulty x Replication $F(3,40) = 4.7$, $p < .05$, and a four-way interaction of Instruction x Appropriateness x Difficulty x Replication $F(3,40) = 3.3$, $p < .05$. These interactions seem to be due to one condition, the subjects in the Eff-Only Replication who studied the Rule Alone instructions. The effects of both Difficulty and Appropriateness were twice as big for this group than any other group. In the Decision interval, there was a main effect of Replication, $F(1,40) = 7.8$, $p < .01$. Subjects in the Eff-Only group were 2.3 seconds faster than subjects in the Eff&Adv group. The Replication factor also interacted with Appropriateness, $F(1,40) = 4.0$, $p = .05$. Subjects in the Eff-Only Replication were equally fast on Appropriate and Inappropriate trials (6.0 and 5.8 secs., respectively), while the Eff&Adv subjects were about a second faster on the Appropriate trials than the Inappropriate trials (7.8 vs. 8.7 secs).

Finally, on the recall measure, subjects in each of the two replications recalled the advice equally well: there was no main effect of Replication and this factor did not interact with any other factor.

These effects and interactions weaken the generality of the findings since they suggest that effects of advice vary for different groups of subjects or for other uncontrolled reasons. In general, the pattern of effects suggests simply that the Eff-Only group was quicker and more accurate than the Eff&Adv subjects. The overall speed and accuracy advantages of the Eff-Only group may be due to the different student profiles at Carnegie-Mellon and the University of Pittsburgh. Carnegie-Mellon students also are more likely to have training in problem solving and other skills that might have been relevant to these tasks. In order to account for these effects, it will be necessary to collect demographic information on subjects in future studies.

including such information as SAT scores, GPA, spatial reasoning ability, major field of study, and so on

Chapter 7
CONCLUSION

7.1. The Role of Elaborations in Instructional Texts

The goal of this research was to find out what makes for a more effective instructional text for skill learning. The traditional wisdom on the content of such texts assumed that learners benefit from detailed explanations of every relevant point. The research of Reder, Charney & Morgan (in press) and Carroll (1985) suggested that this is not the case. In fact, Reder, Charney & Morgan's results suggested that learners benefit only from elaborations on particular types of information in a computer user's manual. Both novice and experienced computer users performed tasks on the computer more efficiently if they had studied elaborations of the command syntax, but elaborations about the function of the commands or when to apply them had no effect on performance.

The present research was designed to explore this result further. Is it true that elaborations on "selection information" have no benefit for learners? Or was the lack of any benefit due to the particular tasks or the particular elaborations we had chosen in the experiment? The answers to these questions are important because of their implications for writing manuals. If writers know in advance that elaborations on selection information don't help learners, but that elaborations on command syntax do, then they can allocate their efforts accordingly. On the other hand, if the

benefit of selection elaborations depends on the task. Then writers must do careful analyses of the kinds of tasks learners might perform in order to decide whether elaborations on the selection information are worthwhile.

In an attempt to find out whether selection information benefits from elaboration, I designed a new task in which careful selection between commands was essential to good performance. Selection information was stated as an advisory rule for choosing between commands. In some manuals, the advice was unelaborated (Rule Alone), while in others it was elaborated with one of two forms of examples (Concept Example or Task Example). Examples were chosen to elaborate the advice because examples have often proven quite effective for learning to apply rules.

The results of the experiment suggest that two groups of subjects, the subjects who studied unelaborated advice (the Rule Alone group) and the subjects who studied advice elaborated with a Concept Example did follow the advice. These subjects relied on the advice equally strongly, and did so regardless of whether tasks were easy or difficult. However, following the advice in this experiment did not improve subjects' performance over subjects who saw no advice at all. The failure of the advice to improve performance suggests either that the advice itself was deficient or that subjects did not learn how to apply the advice well enough. In either case, the experiment sheds some light on the effectiveness of different forms of advice and on how willing people are to follow the advice.

To the extent that elaborating the advice with an example did not improve performance significantly over having the advice alone, the evidence suggests that selection information may really be a type of information that doesn't benefit from elaboration. However, the data are not clean enough to permit such a conclusion.

For one thing, subjects who saw the Task Example did not behave the same way as the rest of the subjects who saw advice. The Task Example subjects may not have paid enough attention to the advice or they may have reinterpreted it in an unexpected way. Since the Task Example was expected to be the most effective form of elaboration, it is important to investigate what happened; perhaps if subjects had been able to make proper use of the Task Example, their performance would have been superior to any other group, including subjects who saw no advice.

The fact that subjects apparently followed the advice even though it did not improve their performance tells us something about how willing people are to rely on advice. We have several other indications in this experiment that people rely heavily on advice. First, subjects were consistently misled by the advice into making wrong answers on the Inappropriate trials. Second, subjects relied on the advice equally heavily on Easy trials as on Hard trials. This degree of reliance on advice is interesting as well as slightly worrisome. It is interesting because teachers have reported needing special efforts to get students to employ general problem-solving heuristics (e.g., Schoenfeld, 1979). Perhaps people were more willing to rely on the advice because it was related directly to a specific procedure (i.e., the Change or Move command) in a specific domain (i.e., Box-World). The reliance on advice is worrisome because we would like students to use advice judiciously. It is possible that with more experience, subjects would come to rely less on the advice. In any case, a lesson that clearly emerges from this experiment is that writers should take care when giving advice. Advice that seems sensible and helpful doesn't always improve performance. Furthermore, writers should try to assess how often advice is inappropriate, since people who follow the advice may easily be led astray.

7.2. Goals for Future Research

The research presented here points the way to future studies on two different levels: the local level of this experimental paradigm and the more general level of the content of instructional texts.

First, the results of the experiment provided a number of surprises and puzzles that bear further investigation. These include:

- the apparent unhelpfulness of the advice in general;
- the unexpected behavior of the Task Example subjects, who behaved unlike the other groups that saw advice;
- the unusual behavior of subjects in the Move Isomorph, whose performance seemed to suffer when they tried to follow the advice; and
- the poor performance of subjects in general on the Recall and Advice Recognition tasks.

Specific proposals for investigating some of these puzzles were presented in Chapters 5 and 6.

On the more general level of research into the content of instruction, there is also much research to be done. First, we obviously can't conclude from this study that selection information is a type of information that needs no elaboration. Logically, we will never be able to demonstrate this fact directly. Whenever a study shows that selection elaborations fail to improve performance, it will be possible to claim that the study still employed the wrong tasks or the wrong elaborations.

There are two ways to break out of this cycle. One way would be to build up a large body of evidence from studies like the present one which span a variety of task domains and use the most plausible tasks and elaborations possible. If we find

the same result over and over again, it might be reasonable to conclude that selection information does not benefit from elaboration. The second way to break the cycle would be to adopt a more theoretical approach. If we assume for the sake of argument that people don't perform better after seeing elaborations about selection information, then it is interesting to ask why they don't and why people do seem to benefit from elaborations about the command syntax. If we can pinpoint the differences in these types of information or how people make use of them, then we may be able to correctly predict what other types of information do and do not benefit from elaboration.

This kind of effort requires identifying significant similarities and differences between information types. The research reported here attempted to exploit one similarity between syntax information and selection information: both types of information can be expressed as rules that state what actions to perform in particular situations. This was why examples were chosen as the form of elaboration for the advice: perhaps examples always help people apply information that comes in the form of a rule. However, the results of this experiment indicate that concept examples did not help subjects apply the advice more effectively. The behavior of the Task Example group merits further investigation on this account.

This research also proposed a difference between syntax information and selection information. Employing the former type of information is compulsory, but the latter is discretionary. That is, people cannot perform any tasks on the computer without knowing how to issue the commands they select correctly. However, people can often function perfectly well without selecting the most appropriate command; they may be less efficient, but they get the job done. Perhaps elaborations are only useful for compulsory information. In any case, a closer study of how subjects went

about solving the problems in the Efficiency task may help shed more light on what kinds of information they needed.

In sum, systematic research on what information learners need and why they need it promises to help us achieve the goal of discovering how to write more effective instructional texts.

Appendix A
BOX-WORLD MANUAL

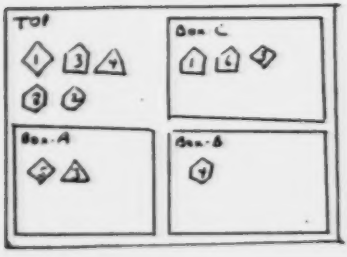
Introduction

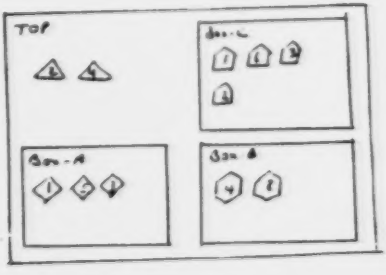
Box-World is a computer program for manipulating simple geometric Objects that are contained in Boxes, as pictured in the diagrams below. This manual will teach you how to write commands to make the computer delete Objects, move them from one box to another or change their shape, so that you can accomplish goals like taking the configuration of Objects pictured below on the left and turning it into the configuration pictured on the right. Some general features of Box-World are described below. On the following pages, you will find instructions for how to change, move and delete Objects.

Types of Objects. The Objects in the Box-World domain consist of five simple polygons: triangles, diamonds, pentagons, hexagons and heptagons.

Arrangements of Boxes. Boxes appear on the screen as rectangular outlines. A Box-World arrangement always starts with one very large box named TOP. TOP can contain up to three Boxes, each of which can contain even smaller boxes.

Identifying Boxes and Objects. In your commands to the computer, you will have to refer to Boxes and Objects by name. Boxes are identified by a name, such as TOP or BOX-A, printed in the upper left-hand corner of the box. Objects have a number (from 0 to 9) printed inside them. For example, a triangle with the number 3 inside it is referred to as TRIANGLE3, in order to distinguish it from DIAMOND3 or TRIANGLE5.





Changing the Shape of an Object

You can use the CHANGE command to change the shape of an Object. For example, you can change a pentagon into a hexagon or a diamond into a triangle. There is a limit on what kind of change you can make with a single change command: you can only add or take away one "side." To see an example, consider the list of possible Box-World Objects arranged below according to how many sides they have. By issuing one CHANGE command, you can change a diamond into either a triangle or a pentagon. To change the diamond into any other shape, you would have to issue a series of CHANGE commands, one for each side you wish to add or take away. So, to change a diamond into a hexagon, you would issue one command to change it into a pentagon and another to change the pentagon into a hexagon. With multiple CHANGE commands, you can change any Object into any of the other four shapes. Obviously, a triangle cannot turn into an Object with two sides and a heptagon cannot turn into an Object with eight sides.

Heptagon (7 sides)

Hexagon (6 sides)

Pentagon (5 sides)

Diamond (4 sides)

Triangle (3 sides)

FORMAT

A CHANGE command has three parts which must be typed in the following order:

CHANGE [NAME OF OBJECT] [TYPE OF CHANGE]

The first part is CHANGE, the name of the command. The second part is the name of the Object, such as PENTAGON1. The third part is the direction of the change: type either MORE-SIDES to get a shape with one more side than the Object has now or FEWER-SIDES to get a shape with one fewer sides. To finish issuing the command to the computer, press the RETURN key.

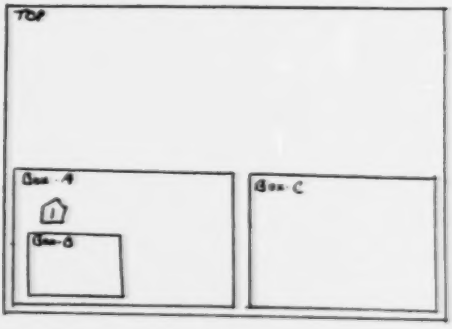
EXAMPLE

The following command takes PENTAGON1 and changes it into a diamond (by making it have FEWER SIDES):

CHANGE PENTAGON1 FEWER-SIDES

Moving an Object to Another Box

You can use the MOVE command to put an Object into a different Box. Individual moves are limited to crossing the edge of only one Box. This means that with a single MOVE command, you can only move an Object into the next larger or next smaller box. For example, suppose the Object you want to move is PENTAGON1 which is inside BOX-A. By issuing one MOVE command, you can either move the Object down into BOX-B, a smaller box inside BOX-A, or move it up into TOP which contains BOX-A. To move PENTAGON1 into any other box, you would have to issue a series of MOVE commands, one for each box whose edge you would have to cross. So, to move PENTAGON1 to BOX-C, you would issue one command to move it to TOP and another to move it from TOP to BOX-C. With multiple MOVE commands, you can move any Object into any box on the screen.



FORMAT

A MOVE command has three parts which must be typed in the following order:

MOVE [NAME OF OBJECT] [DESTINATION]

The first part is MOVE, the name of the command. The second part is the name of the Object, such as PENTAGON1. The third part is the name of the box that you want to move the Object to, such as BOX-B. To finish issuing the command to the computer, press the RETURN key. If you typed a destination that is too "far away" (that is, you would have to cross the edge of more than one box to get there), the computer will respond "Destination too far" and you will have to issue a new MOVE command.

EXAMPLE

The following command takes PENTAGON1 and puts it into BOX-B:

MOVE PENTAGON1 BOX-B

Deleting an Object

You can erase an Object from the screen using the DELETE command. Once you delete an Object, you cannot bring it back.

FORMAT

The DELETE command has two parts which must be supplied in the following order:

DELETE [NAME OF OBJECT]

The first part is DELETE, the name of the command. The second part is the name of the Object you want to delete, such as PENTAGON1.

EXAMPLE

The following command erases PENTAGON1 from the Box-World:

DELETE PENTAGON1

Appendix B
BOX-WORLD TASKS

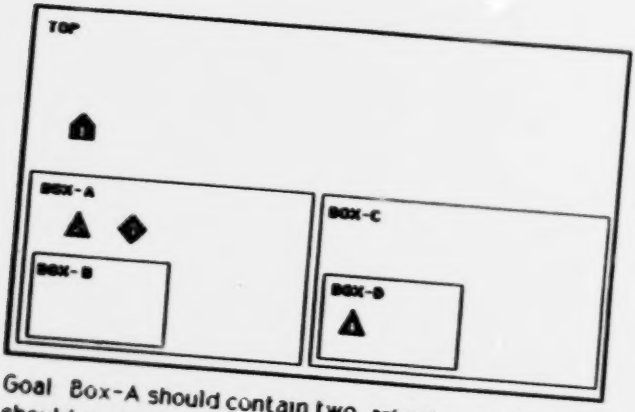
Box-World Trial (B-4)
Change Isomorph
Easy, Appropriate, Use

B-4 is a Use problem because you want two triangles in Box-A, and there's a diamond in Box-A that you can change into a triangle with one step. The correct response is No because the action proposes a different way to solve the problem, moving in the triangle from Box-D.

Box-World Trial (B-4)

Change Isomorph

Easy, Appropriate, Use



Goal Box-A should contain two triangles only, and Box-B should contain one pentagon only

Action Move triangle-4 to Box-A

Most efficient solution

Change diamond-3 into a triangle -- 1 step
Move pentagon-1 to Box-B -- 2 steps

Efficiency Task response No

Advice Recognition response No

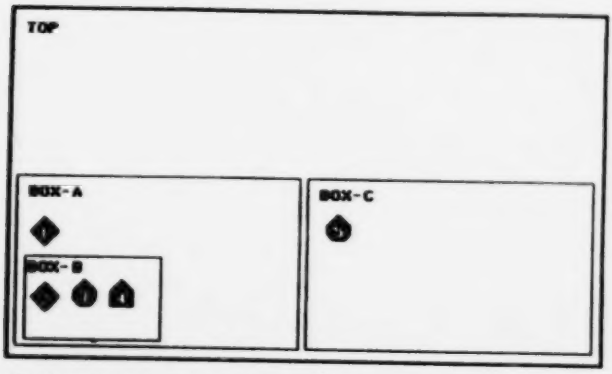
Box-World Trial (C-7)
Change Isomorph
Easy, Inappropriate, Don't Use

C-7 is a Don't Use problem because you want another heptagon in Box-B, but the advice says not to try to achieve that goal by changing the pentagon. In this case, the advice is inappropriate, because changing the pentagon to a heptagon is part of the most efficient solution.

Box-World Trial (C-7)

Change Isomorph

Easy, Inappropriate, Don't Use



Goal: Box-B should contain two diamonds and two heptagons only

Action: Change pentagon-4 into a heptagon

Most efficient solution

Change pentagon-4 into a heptagon	2 steps
Move diamond-1 to Box-B	1 step

Efficiency Task response: Yes

Advice Recognition response: No

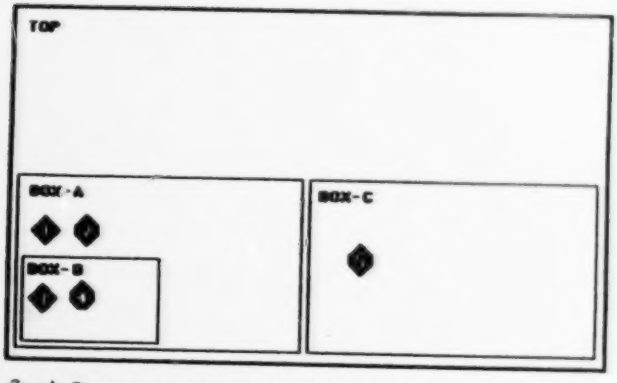
Box-World Trial (E-6)
Change Isomorph
Hard, Appropriate, Don't Use

E-6 is a Don't Use problem because you want another diamond in Box-A and the advice says not to try to achieve that goal by changing the hexagon. In this case, the advice is appropriate. The most efficient overall solution solves that subgoal by moving in the diamond from Box-B.

Box-World Trial (E-6)

Change Isomorph

Hard, Appropriate, Don't Use



Goal: Box-A should contain two diamonds only, and Box-C should contain one pentagon and one heptagon only.

Action: Change hexagon-2 into a diamond.

Most efficient solution:

Move diamond-3 to Box-A	1 step
Move hexagon-2 to Box-C	2 steps
Change hexagon-2 into a pentagon	1 step
Change hexagon-5 into a heptagon	1 step

Efficiency Task response: No

Advice Recognition response: No

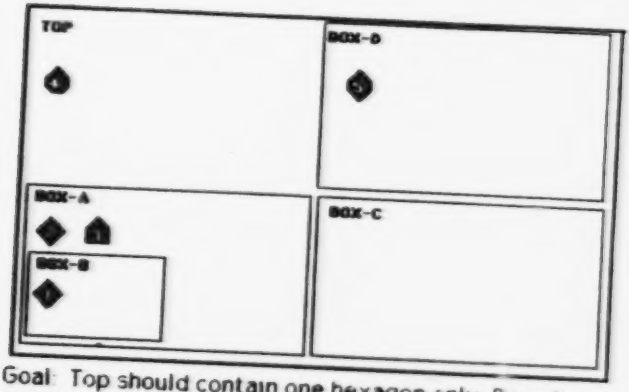
Box-World Trial (E-18)
Change lexmargh
Hard, Inappropriate, Use

E-18 is a Use problem because you want another diamond in Box-A, and there is a pentagon in Box-A that can be changed into a diamond in just one step. In this case, the advice is inappropriate – it's more efficient to move in the diamond from Box-B.

Box-World Trial (E-15)

Change Isomorph

Hard, Inappropriate, Use



Goal: Top should contain one hexagon only, Box-A should contain two diamonds only and Box-C should contain one pentagon and one heptagon only

Action: Move pentagon-3 to Box-C

Most efficient solution:

Move pentagon-3 to Box-C	2 steps
Move diamond-1 to Box-A	1 step
Move heptagon-4 to Box-C	1 step
Move hexagon-5 to Top	1 step

Efficiency Task response: Yes

Advice Recognition response: No

Appendix C

BOX-WORLD TASKS – MOVE IS

MORPHS

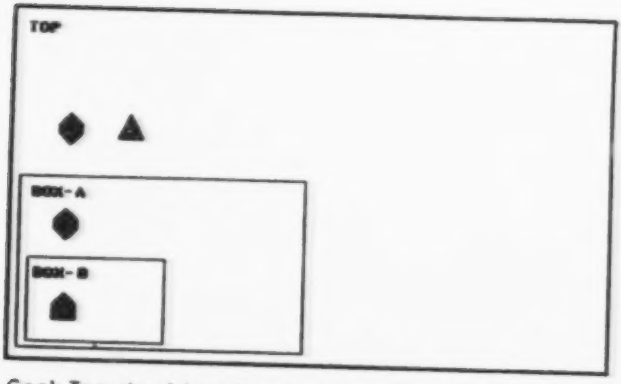
Box-World Trial (B-4)
Move Isomorph
Easy, Appropriate, Use

B-4 is a Use problem because you want two hexagons in Top, and there's a hexagon in Box-A that you can move to Top with one step. The correct response is No because the action proposes a different way to solve the problem, changing the triangle in Top to a hexagon.

Box-World Trial (B-4)

Move Isomorph

Easy, Appropriate, Use



Goal: Top should contain two hexagons, among other objects, Box-A should be empty, and Box-B should contain one heptagon, among other objects

Action: Change triangle-4 into a hexagon

Most efficient solution:

Move hexagon-3 to Top -- 1 step

Change pentagon-1 into a heptagon -- 2 steps

Efficiency Task response: No

Advice Recognition response: No

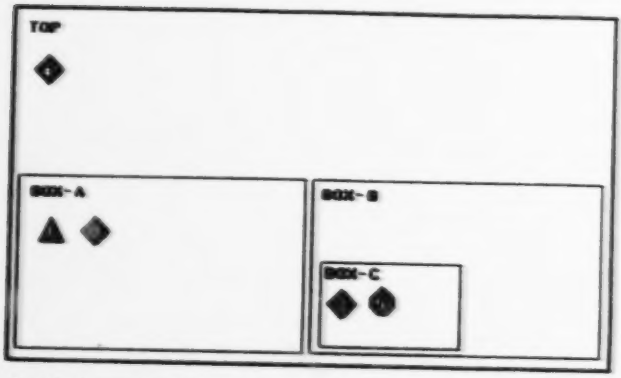
Box-World Trial (C-7)
Move Isomorph
Easy, Inappropriate, Don't Use

C-7 is a Don't Use problem because you want another diamond in Box-C, but the advice says not to try to achieve that goal by moving the diamond from Top. In this case, the advice is inappropriate, because moving the diamond from Top is part of the most efficient solution.

Box-World Trial (C-7)

Move Isomorph

Easy, Inappropriate, Don't Use



Goal: Top should be empty, Box-A should contain two diamonds among other objects, and Box-C should contain two diamonds among other objects

Action: Move diamond-4 to Box-C

Most efficient solution:

Move diamond-4 to Box-C	2 steps
Change triangle-1 into a diamond	1 step

Efficiency Task response: Yes

Advice Recognition response: No

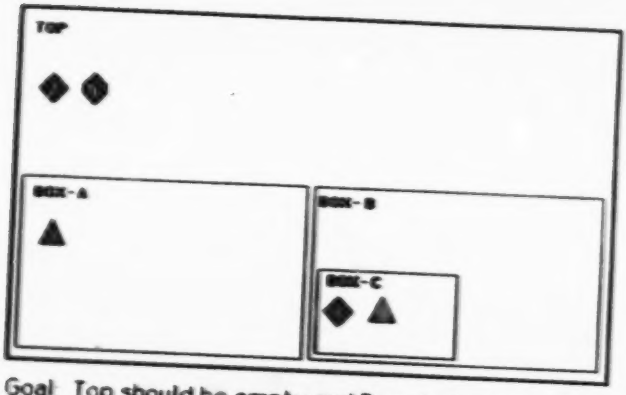
Box-World Trial (E-6)
Move Isomorph
Hard, Appropriate, Don't Use

E-6 is a Don't Use problem because you want another diamond in Box-C and the advice says not to try to achieve that goal by moving in diamond-2 from Top. In this case, the advice is appropriate. The most efficient overall solution solves that subgoal by changing triangle-3 into a diamond.

Box-World Trial (E-6)

Move Isomorph

Hard, Appropriate, Don't Use



Goal: Top should be empty and Box-A should contain one hexagon among other objects, Box-B should contain one hexagon among other objects and Box-C should contain two diamonds among other objects

Action: Move diamond-2 to Box-C

Most efficient solution:

Change triangle-3 into a diamond	1 step
Change diamond-3 into a hexagon	2 steps
Move hexagon-2 to Box-B	1 step
Move hexagon-5 to Box-A	1 step

Efficiency Task response: No

Advice Recognition response: No

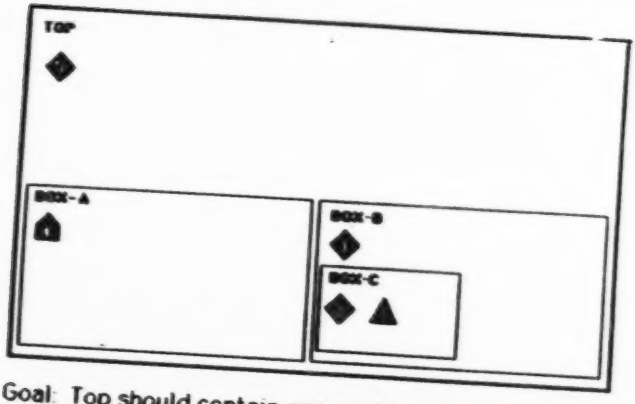
Box-World Trial (E-15)
Move Isomorph
Hard, Inappropriate, Use

E-15 is a Use problem because you want another diamond in Box-C, and there is a diamond in Box-B that can be moved to Box-C with just one step. In this case, the advice is inappropriate – it's more efficient to change triangle-1 into a diamond.

Box-World Trial (E-15)

Move Isomorph

Hard, Inappropriate, Use



Goal: Top should contain one pentagon among other objects, Box-A should contain one hexagon among other objects, Box-B should contain one hexagon among other objects and Box-C should contain two diamonds among other objects

Action: Change diamond-3 into a hexagon

Most efficient solution:

Change diamond-3 into a hexagon	2 steps
Change triangle-1 into a diamond	1 step
Change pentagon-4 into a hexagon	1 step
Change diamond-5 into a pentagon	1 step

Efficiency Task response: Yes
Advice Recognition response: No

Appendix D

DISTRIBUTION OF SUBJECTS AMONG CONDITIONS

APPENDIX D

Distribution of Subjects among Conditions for
Recall of Advice, Advice Recognition and Efficiency Tasks

	CHANGE ISOMORPH				MOVE ISOMORPH			
	No Advice	Rule Alone	Concpt Eg	Task Eg	No Advice	Rule Alone	Concpt Eg	Task Eg
RECALL OF ADVICE	0 (0)	11 (5)	12 (6)	11 (5)	0 (0)	17 (6)	17 (6)	15 (5)
ADVICE RECOGNITION TASK	0 (0)	6 (0)	6 (0)	6 (0)	0 (0)	5 (0)	6 (0)	6 (0)
EFFICIENCY TASK	12 (6)	12 (6)	12 (6)	12 (6)	15 (6)	17 (6)	17 (6)	16 (6)

The figures in parentheses represent the number of these subjects who were in the Eff-Only Replication group.

Appendix E

SAMPLE RECALL RESPONSES

Sample Responses for Recall of Advice

Responses scored as 0:

- Don't remember.
- I do not recall that the manual explained when to use commands. They only said what the commands did and their syntax.
- You can move objects one box at a time. You can change objects by adding or deleting sides one at a time. You can delete objects; once deleted, they can't return.
- Use commands when they lead to the most efficient way of achieving the goal.
- Count the number of steps involved in selecting one strategy and compare that with the number of steps involved in selecting the other. Choose the strategy that would lead you to the goal with the least number of steps.

Responses scored as .5:

- Change things first but don't move them much.
- The manual said that it was better not to change sides unless necessary because it took up too many moves.
- Use "move" only if it results in less steps than changing the number of sides.
- Use the change command as little as possible, i.e., only when it would take fewer moves to change an object than to move it.
- Use change command only when it is necessary to use the least number of commands.

Responses scored as 1.0:

* It suggested to use the change command when a object needs to be changed by only one side. And to move objects from box to box when necessary.

* If the number of sides to be increased or decreased is more than one, find another way of solving the problem.

* If you have to move thru more than one box, don't use move command.

* Use the "move" command when only one move was necessary.

* The advice said not to use the CHANGE command to go from one shape to another which was several steps away. It said that a more efficient way could usually be found.

REFERENCES

Allwood, C. M., Wikstrom, T., & Reder, L. M. (1982). Effects of presentation format on reading retention: Superiority of summaries in free recall. *Poetics, 11*, 145-153.

Anderson, J. R. (1974a). Retrieval of propositional information from long-term memory. *Cognitive Psychology, 5*, 451-474.

Anderson, J. R. (1983). *The Architecture of Cognition*. Cambridge, MA: Harvard University Press.

Anonymous. (1983). *Personal Computer Computer Language Series. Disk Operating System*. Boca Raton, FL: IBM. Version 2.00.

Atlas, M. (1980). *The User Edit: Making Manuals Easier to Use* Technical Report IBM.

Bransford, J. (1979). *Human Cognition*. Belmont, CA: Wadsworth.

Bugelski, B. R. (1962). Presentation time, total time, and mediation in paired associate learning. *Journal of Experimental Psychology, 63*, 409-412.

Card, S., Moran, T., & Newell, A. (1983). *The Psychology of Human Computer Interaction*. Hillsdale, NJ: Erlbaum.

Carroll, J. (March 1985). *Designing MINIMALIST Training Materials* Research Report 48843 IBM Watson Research Center, Computer Science Department. Also appeared in Datamation 30(18), 125-136, 1984.

Charney, D. (1983). The Role of Examples and Details in the Comprehension of Text. Unpublished mss., Carnegie-Mellon University.

Charney, D. (1984b). Redesigning and Testing a Work Order Form. *Information Design Journal, 4*(2), 131-148.

Charney, D. (1984a). The Validity of Using Holistic Scoring to Evaluate Writing. *Research in the Teaching of English, 18*(1), 65-81.

Cooper, E. H., & Pantle, A. J. (1967). The total-time hypothesis in verbal learning. *Psychological Bulletin, 68*, 221-234.

Duffy, T., & Kabance, P. (1982). Testing a Readable Writing Approach to Text Revision. *Journal of Experimental Psychology, 5*, 733-748.

Elio, R. & Anderson, J. R. (1984). The Effects of Information Order and Learning Mode on Schema Abstraction. *Memory & Cognition, 12*(1), 20-30.

Felker, D. (April 1980). *Document Design: A Review of the Relevant Research*. Technical Report. Washington, D.C.: American Institute for Research.

Gilson, C., and Abelson, R. (1968). The Subjective Use of Inductive Evidence. In P. C. Wason & P. N. Johnson-Laird (Ed.), *Thinking & Reasoning*. Harmondsworth, Middlesex, England: Penguin Books.

Hayes, J.R., & Simon, H. (1976). Psychological Differences among Problem Isomorphs. In H. Simon (Ed.), *Models of Thought*. New Haven: Yale University Press. Reprinted from N.J. Castellan, D.B. Pisoni & G.R. Potts (eds.), *Cognitive Theory*, Vol. 2, pp. 21-41, Potomac, MD: Erlbaum, 1977.

Hintzman, D. (1980). Differential Forgetting of Prototypes and Old Instances. *Memory and Cognition*, *8*(4), 378-382.

Hobbs, J. (November 1978). *Why is Discourse Coherent?* Technical Note. SRI International.

Holland, M., Rose, A., Dean, R., & Dory, S. (February 1985). *Processes Involved in Writing Effective Procedural Instructions*. Final Report. American Institute for Research, Washington, D.C.

Kieras, D. (February 1985). *The Role of Prior Knowledge in Operating Equipment from Written Instructions*. Technical Report. University of Michigan.

Kieras, D. (June 1985). *Improving the Comprehensibility of a Simulated Technical Manual*. Technical Report. University of Michigan.

Kinneavy, J. (1971). *A Theory of Discourse*. Englewood Cliffs, NJ: Prentice-Hall.

LeFevre, J. (Spring 1985). *A Model of the Use of Instruction and Example Information on a Simple Inductive-Reasoning Task*. Master's thesis, Department of Psychology, University of Alberta.

LeFevre, J. & Dixon, P. (1984). Do Written Instructions Need Examples. Unpublished mss., University Alberta.

Mandl, H., Schnotz, W., & Tergan, S. (1984). On the function of examples in instructional texts. Paper presented at the AERA meeting, New Orleans, April, 1984.

Merrill, M. D. (1983). Component Display Theory. In Charles Reigeluth (Ed.), *Instructional-Design Theories and Models: An overview of their current status*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.

Nitsch, K. E. (1977). Structuring Decontextualized Forms of Knowledge. Unpublished doctoral dissertation, Vanderbilt University.

Paivio, A., Yuille, J. C., & Madigan, S. A. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology Monograph Supplement*, *76*(1), Pt. 2.

Pepper, J. (1981). Following Students' Suggestions for Rewriting Computer Programming Textbooks. *American Education Research Journal, 18*, 259-269.

Perelman, C., & Olbrechts-Tyteca, L. (1969). *The New Rhetoric: A Treatise on Argumentation*. Notre Dame, Indiana: University of Notre Dame Press.

Pirolli, P. & Anderson, J. R. (1985). The Role of Learning from Examples in the Acquisition of Recursive Programming Skills. *Canadian Journal of Psychology, 39*, 240-272.

Price, J. (1984). *How to Write a Computer Manual: A Handbook of Software Documentation*. Menlo Park, CA: The Benjamin/Cummings Publishing Co.

Reder, L. M. (1982). Elaborations: When do they help and when do they hurt? *Text, 2*, 211-224.

Reder, L. M., & Anderson, J. R. (1980a). A comparison of texts and their summaries: Memorial consequences. *Journal of Verbal Learning and Verbal Behavior, 19*, 121-134.

Reder, L. M., & Anderson, J. R. (1982). Effects of spacing and embellishment on memory for the main points of a text. *Memory and Cognition, 10*, 97-102.

Reder, L., Charney, D., & Morgan, K. (August 1984). *The Role of Elaborations in Learning a Skill from an Instructional Text* Technical Report 1 Carnegie-Mellon University.

Reder, L., Charney, D., & Morgan, K. (In press). The Role of Elaborations in Learning a Skill from an Instructional Text. *Memory and Cognition*, . In press.

Reder, L. (). Beyond Associations: Strategic Components in Memory Retrieval. To appear in D. Gorlich & R. Hoffman, (Eds.), Learning and Memory: The Ebbinghaus Centennial Conference. Hillsdale, NJ: Lawrence Erlbaum & Associates.

Reigeluth, C., & Stein, F. (1983). The Elaboration Theory of Instruction. In Charles Reigeluth (Ed.), *Instructional-Design Theories and Models: An overview of their current status*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Riley, M. & O'Malley, C. (March 1984). *Planning Nets: A Framework for Analyzing User-Computer Interactions* ICS Report UCSD.

Rosch, E. (1977). Classification of real-world objects: Origins and representations in cognition. In P.N. Johnson-Laird & P.C. Wason (Ed.), *Thinking*. Cambridge: Cambridge University Press.

Ross, B. (1984). Remindings and their effects in learning a cognitive skill. *Cognitive Psychology, 16*, 371-416.

Scharer, L. (July 1983). User Training: Less is More. *Datamation, 29*, 175-182.

Schoenfeld, A. (1979). Can Heuristics Be Taught? In J. Lockhead & J. Clements (Ed.), *Cognitive Process Instruction*. Philadelphia: The Franklin Institute Press.

Sebrechts, M., Galambos, J., Wagner, R., Black, J., Deck, J., & Wilder, E. (1983).

The Effects of Diagrams on Learning to Use a System. To appear as Learning and Using Systems Report No. 2.

Smedslund, J. (1968). Mental Processes Involved in Rapid Logical Reasoning. *Scandanavian Journal of Psychology, 9,* 187-205.

Smith, E., & Goodman, L. (July 1982). *Understanding Instructions: The Role of Explanatory Material* Technical Report Bolt, Beranek and Newman Inc.

Sternberg, R. & Ketron, J. (1982). Selection and Implementation of Strategies in Reasoning by Analogy. *Journal of Educational Psychology, 74*(3), 399-413.

Sullivan, P., & Flower, L. (forthcoming). How do Users Read Computer Manuals? Some Protocol Contributions to Writers' Knowledge. In Bruce T. Peterson (Ed.), *Convergences: Essays on Reading, Writing, and Literacy.* Urbana, IL: National Council of Teachers of English.

Swaney, J., Janik, C., Bond, S., & Hayes, J. R. (June 1981). *Editing for Comprehension: Improving the Process through Reading Protocols* Communications Design Center Technical Report Carnegie-Mellon University.

Tausworthe, R. (1979). *Standardized Development of Computer Software, Part II.* Englewood Cliffs, NJ: Prentice-Hall.

Tennyson, R. (1973). Effect of Negative Instances in Concept Acquisition Using a Verbal Learning Task. *Journal of Experimental Psychology, 84,* 247-260.

Tennyson, R., Steve, M., & Boutwell, R. (1975). Instance sequence in concept acquisition. *Journal of Experimental Psychology, 87,* 821-827.

Tennyson, R., Woolley, F., & Merrill, M. (1972). Exemplar and Nonexemplar Variables which Produce Correct Concept Classification Behavior and Specified Classification Errors. *Journal of Educational Psychology, 63,* 144-152.

Williams, J. (1981). *Style: Ten lessons in clarity and grace.* Glenview, IL: Scott, Foresman and Company.

Wright, P. (1977). Presenting Technical Information: A survey of research findings. *Instructional Science, 6,* 93-134.

Wright, P. (1978). Feeding the Information Eaters: Suggestions for Integrating pure and applied research on language comprehension. *Instructional Science, 7,* 249-312.

Wright, P. (December 1983). Manual Dexterity: A User-Oriented Approach to Creating Computer Documentation. In A. Janda (Ed.), *Human Factors in Computing Systems.* Boston: Computer-Human Interaction.

Young, R., Becker, A., Pike, K. (1970). *Rhetoric: Discovery and Change.* New York: Harcourt, Brace and World.

Dr. Robert Ahlers
Code N711
Human Factors Laboratory
NAVTRAEQUIPCEN
Orlando, FL 32813

Dr. Ed Aiken
Navy Personnel R&D Center
San Diego, CA 92152

Dr. Earl A. Alluisi
HQ, AFHRL (AFSC)
Brooks AFB, TX 78235

Technical Director, ARI
5001 Eisenhower Avenue
Alexandria, VA 22333

Dr. Meryl S. Baker
Navy Personnel R&D Center
San Diego, CA 92152

Dr. Thomas G. Bever
Dept. of Psychology
Columbia University
New York, NY 10027

Dr. Menucha Birenbaum
School of Education
Tel Aviv University
Tel Aviv, Ramat Aviv 69978
ISRAEL

Dr. Gautam Biswas
Department of Computer Science
University of South Carolina
Columbia, SC 29208

Dr. John Black
Yale University
Box 11A, Yale Station
New Haven, CT 06520

Dr. Arthur S. Blaiwes
Code N711
Naval Training Equipment Center
Orlando, FL 32813

Dr. Robert Blanchard
Navy Personnel R&D Center
San Diego, CA 92152

Dr. Richard Braby
NTSC Code 10
Orlando, FL 32751

Dr. Robert Breaux
Code N-095R
NAVTRAEQUIPCEN
Orlando, FL 32813

Dr. Ann Brown
Center for the Study of Reading
University of Illinois
51 Gerty Drive
Champaign, IL 61280

Dr. Patricia A. Butler
NIE Mail Stop 1806
1200 19th St., NW
Washington, DC 20208

Dr. Robert Calfee
School of Education
Stanford University
Stanford, CA 94305

Dr. Pat Carpenter
Carnegie-Mellon University
Department of Psychology
Pittsburgh, PA 15213

Dr. Robert Carroll
NAVOP 01B7
Washington, DC 20370

Dr. Fred Chang
Navy Personnel R&D Center
Code 51
San Diego, CA 92152

Dr. Davida Charney
Department of Psychology
Carnegie-Mellon University
Schenley Park
Pittsburgh, PA 15213

Dr. Eugene Charniak
Brown University
Computer Science Department
Providence, RI 02912

Mr. Raymond E. Christal
AFHRL/MOE
Brooks AFB, TX 78235

Director,
Manpower Support and
Readiness Program
Center for Naval Analysis
2000 North Beauregard Street
Alexandria, VA 22311

Chief of Naval Education
and Training
Liaison Office
Air Force Human Resource Laboratory
Operations Training Division
Williams AFB, AZ 85224

Assistant Chief of Staff
for Research, Development,
Test, and Evaluation
Naval Education and
Training Command (N-5)
NAS Pensacola, FL 32508

Dr. John J. Collins
Director, Field Research
Office, Orlando
NPRDC Liaison Officer
NTSC Orlando, FL 32813

Dr. Stanley Collyer
Office of Naval Technology
Code 222
800 N. Quincy Street
Arlington, VA 22217-5000

LT Judy Crookshanks
Chief of Naval Operations
OP-112G5
Washington, DC 20370-2000

Dr. Mary Cross
Department of Education
Adult Literacy Initiative
Room 4145
400 Maryland Avenue, SW
Washington, DC 20202

CTB/McGraw Hill Library
2500 Garden Road
Monterey, CA 93940

CAPT P. Michael Curran
Office of Naval Research
800 N. Quincy St.
Code 125
Arlington, VA 22217-5000

Dr. Cary Czichon
Mail Station 3407
Texas Instruments AI Lab
P.O. Box 405
Lewisville, TX 75067

Bryan Dallman
AFHRL/LRT
Lowry AFB, CO 80230

Dr. Natalie Dehn
Department of Computer and
Information Science
University of Oregon
Eugene, OR 97403

Coery Delacote
Directeur de L'informatique
Scientifique et Technique
CNRS
15, Quai Anatole France
75700 Paris FRANCE

Dr. Sharon Derry
Florida State University
Department of Psychology
Tallahassee, FL 32306

Defense Technical
Information Center
Cameron Station, Bldg 5
Alexandria, VA 22314
Attn: TC
(12 Copies)

Dr. Thomas M. Duffy
Communications Design Center
Carnegie-Mellon University
Schenley Park
Pittsburgh, PA 15213

Dr. Richard Duran
University of California
Santa Barbara, CA 93106

Barbara Eason
Military Educator's
Resource Network
InterAmerica Research Associates
1555 Wilson Blvd
Arlington, VA 22209

Edward E. Eddowes
CNATRA N301
Naval Air Station
Corpus Christi, TX 78419

Dr. John Ellis
Navy Personnel R&D Center
San Diego, CA 92252

Dr. Jeffrey Elman
University of California,
San Diego
Department of Linguistics, C-008
La Jolla, CA 92093

Dr. Richard Elster
Deputy Assistant Secretary
of the Navy (Manpower)
OASN (M&RA)
Department of the Navy
Washington, DC 20350-1000

Dr. Susan Embretson
University of Kansas
Psychology Department
Lawrence, KS 66045

Dr. Randy Engle
Department of Psychology
University of South Carolina
Columbia, SC 29208

Dr. William Epstein
University of Wisconsin
W. J. Brogden Psychology Bldg.
1202 W. Johnson Street
Madison, WI 53706

ERIC Facility-Acquisitions
4833 Rugby Avenue
Bethesda, MD 20014

Dr. Edward Esty
Department of Education, OERI
Room 717D
1200 19th St., NW
Washington, DC 20208

Dr. Paul Feltovich
Southern Illinois University
School of Medicine
Medical Education Department
P.O. Box 3926
Springfield, IL 62708

Mr. Wallace Feurzeig
Educational Technology
Bolt Beranek & Newman
10 Moulton St.
Cambridge, MA 02238

Dr. Gerhard Fischer
University of Colorado
Department of Computer Science
Boulder, CO 80309

Dr. Linda Flower
Carnegie-Mellon University
Department of English
Pittsburgh, PA 15213

Dr. Carl H. Frederiksen
McGill University
3700 McTavish Street
Montreal, Quebec H3A 1Y2
CANADA

Dr. John R. Frederiksen
Bolt Beranek & Newman
50 Moulton Street
Cambridge, MA 02138

Dr. R. Edward Geiselman
Department of Psychology
University of California
Los Angeles, CA 90024

Dr. Arthur M. Glenberg
University of Wisconsin
W. J. Brogden Psychology Bldg.
1202 W. Johnson Street
Madison, WI 53706

Dr. Sam Glucksberg
Princeton University
Department of Psychology
Green Hall
Princeton, NJ 08540

Dr. Joseph Goguen
Computer Science Laboratory
SRI International
333 Ravenswood Avenue
Menlo Park, CA 94025

Dr. Susan Goldman
University of California
Santa Barbara, CA 93106

Dr. Sherrie Gott
AFHRL/MODJ
Brooks AFB, TX 78235

Dr. Richard H. Granger
Department of Computer Science
University of California, Irvine
Irvine, CA 92717

Dr. Wayne Gray
Army Research Institute
5001 Eisenhover Avenue
Alexandria, VA 22333

Dr. Henry H. Halff
Halff Resources, Inc.
4918 33rd Road, North
Arlington, VA 22207

Dr. Nancy F. Halff
Halff Resources, Inc.
4918 33rd Road, North
Arlington, VA 22207

Dr. Ronald K. Hambleton
Prof. of Education & Psychology
University of Massachusetts
at Amherst
Hills House
Amherst, MA 01003

Dr. Cheryl Hamel
NTEC
Orlando, FL 32813

Dr. Ray Hannapel
Scientific and Engineering
Personnel and Education
National Science Foundation
Washington, DC 20550

Ms. Carol S. Hargan
Manager
Instructional Videodisc Group
HumRRO
1100 S. Washington
Alexandria, VA 22314

Mr. William Hartung
PEAM Product Manager
Army Research Institute
5001 Eisenhover Avenue
Alexandria, VA 22333

Dr. Wayne Harvey
SRI International
333 Ravenswood Ave.
Room B-S324
Menlo Park, CA 94025

Prof. John R. Hayes
Carnegie-Mellon University
Department of Psychology
Schenley Park
Pittsburgh, PA 15213

Dr. Joan I. Heller
505 Haddon Road
Oakland, CA 94606

Dr. Melissa Holland
Army Research Institute for the
Behavioral and Social Sciences
5001 Eisenhover Avenue
Alexandria, VA 22333

Dr. Claude Janvier
Directeur, CIRADE
Universite' du Quebec a Montreal
Montreal, Quebec H3C 3P8
CANADA

COL Dennis W. Jarvi
Commander
AFHRL
Brooks AFB, TX 78235-5601

Margaret Jerome
c/o Dr. Peter Chandler
83, The Drive
Hove
Sussex
UNITED KINGDOM

Dr. Joseph E. Johnson
Assistant Dean for
Graduate Studies
College of Science and Mathematics
University of South Carolina
Columbia, SC 29208

Col. Dominique Jouslin de Noray
Etat-Major de l'Armee de Terre
Centre de Relations Humaines
3 Avenue Octave Greard
75007 Paris
FRANCE

Dr. Marcel Just
Carnegie-Mellon University
Department of Psychology
Schenley Park
Pittsburgh, PA 15213

Dr. Richard Kern
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Dr. David Kieras
University of Michigan
Technical Communication
College of Engineering
1223 E. Engineering Building
Ann Arbor, MI 48109

Dr. Peter Kincaid
Training Analysis
& Evaluation Group
Department of the Navy
Orlando, FL 32813

Dr. Walter Kintsch
Department of Psychology
University of Colorado
Campus Box 345
Boulder, CO 80302

Dr. David Klahr
Carnegie-Mellon University
Department of Psychology
Schenley Park
Pittsburgh, PA 15213

Dr. Mazie Knerr
Program Manager
Training Research Division
HumRRO
1100 S. Washington
Alexandria, VA 22314

Dr. Janet L. Kolodner
Georgia Institute of Technology
School of Information
& Computer Science
Atlanta, GA 30332

M. Diane Langston
Communications Design Center
Carnegie-Mellon University
Schenley Park
Pittsburgh, PA 15213

Dr. Robert Lawler
Information Sciences, FRL
GTE Laboratories, Inc.
40 Sylvan Road
Waltham, MA 02254

Dr. Alan M. Lesgold
Learning R&D Center
University of Pittsburgh
Pittsburgh, PA 15260

Dr. Alan Leshner
Deputy Division Director
Behavioral and Neural Sciences
National Science Foundation
1800 G Street
Washington, DC 20550

Dr. Clayton Lewis
University of Colorado
Department of Computer Science
Campus Box 430
Boulder, CO 80309

Dr. Charlotte Linde
SRI International
333 Ravenswood Avenue
Menlo Park, CA 94025

Dr. Marcia C. Linn
Lawrence Hall of Science
University of California
Berkeley, CA 94720

Dr. William L. Maloy
Chief of Naval Education
and Training
Naval Air Station
Pensacola, FL 32508

Dr. Manion M. Matthews
Department of Computer Science
University of South Carolina
Columbia, SC 29208

Dr. Richard E. Mayer
Department of Psychology
University of California
Santa Barbara, CA 93106

Dr. James McBride
Psychological Corporation
c/o Harcourt, Brace,
Javanovich Inc.
1250 West 6th Street
San Diego, CA 92101

Dr. Joe McLachlan
Navy Personnel R&D Center
San Diego, CA 92152

Dr. James McMichael
Assistant for MPT Research,
Development, and Studies
NAVOP 01B7
Washington, DC 20370

Dr. Barbara Means
Human Resources
Research Organization
1100 South Washington
Alexandria, VA 22314

Prof. Jacques Mehler
Laboratoire de
Psychologie
54, bvd. Raspail
F-75006 Paris
FRANCE

Dr. Arthur Melmed
U. S. Department of Education
724 Brown
Washington, DC 20208

Dr. George A. Miller
Department of Psychology
Green Hall
Princeton University
Princeton, NJ 08540

Dr. Lance A. Miller
IBM Thomas J. Watson
Research Center
P.O. Box 218
Yorktown Heights, NY 10598

Dr. Lynn Misselt
HQM-222
Control Data Corporation
Box O
Minneapolis, MN 55440

Dr. Andrew R. Molnar
Scientific and Engineering
Personnel and Education
National Science Foundation
Washington, DC 20550

Dr. William Montague
NPRDC Code 13
San Diego, CA 92152

Dr. Tom Moran
Xerox PARC
3333 Coyote Hill Road
Palo Alto, CA 94304

Dr. John Morton
MRC Cognitive
Development Unit
17 Gordon Street
London WC1H OAH
UNITED KINGDOM

Dr. Allen Munro
Behavioral Technology
Laboratories - USC
1845 S. Elena Ave., 4th Floor
Redondo Beach, CA 90277

Spec. Asst. for Research, Experimental & Academic Programs,
NTTC (Code 016)
NAS Memphis (75)
Millington, TN 38054

Assistant for MPT Research,
Development and Studies
NAVOP 0187
Washington, DC 20370

Assistant for Personnel
Logistics Planning,
NAVOP 987H
5D772, The Pentagon
Washington, DC 20350

Dr. Donald A. Norman
Institute for Cognitive Science
University of California
La Jolla, CA 92093

Director, Training Laboratory,
NPRDC (Code 05)
San Diego, CA 92152

Director, Manpower and Personnel
Laboratory,
NPRDC (Code 06)
San Diego, CA 92152

Director, Human Factors
& Organizational Systems Lab,
NPRDC (Code 07)
San Diego, CA 92152

Library, NPRDC
Code P201L
San Diego, CA 92152

Commanding Officer,
Naval Research Laboratory
Code 2627
Washington, DC 20390

Dr. Harry F. O'Neil, Jr.
University of Southern California
School of Education -- VPH 801
Dept. of Educational
Psychology and Technology
Los Angeles, CA 90089-0031

Office of Naval Research,
Code 1142EP
800 N. Quincy Street
Arlington, VA 22217-5000

Office of Naval Research,
Code 1142PT
800 N. Quincy Street
Arlington, VA 22217-5000
(6 Copies)

Psychologist
Office of Naval Research
Branch Office, London
Box 39
FPO New York, NY 09510

Special Assistant for Marine
Corps Matters,
ONR Code 00MC
800 N. Quincy St.
Arlington, VA 22217-5000

Psychologist
Office of Naval Research
Liaison Office, Far East
APO San Francisco, CA 96503

Dr. Judith Orasanu
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Dr. Jesse Orlansky
Institute for Defense Analyses
1801 N. Beauregard St.
Alexandria, VA 22311

CDR R. T. Parlette
Chief of Naval Operations
OP-112G
Washington, DC 20370-2000

Dr. Douglas Pearse
DCIEM
Box 2000
Downsview, Ontario
CANADA

Dr. Nancy Pennington
University of Chicago
Graduate School of Business
1101 E. 58th St.
Chicago, IL 60637

Military Assistant for Training and
Personnel Technology,
OUSD (R & E)
Room 3D129, The Pentagon
Washington, DC 20301

Dr. Ray Perez
ARI (PERI-II)
5001 Eisenhover Avenue
Alexandria, VA 2233

Administrative Sciences Department,
Naval Postgraduate School
Monterey, CA 93940

Dr. Steven Pinker
Department of Psychology
E10-018
M.I.T.
Cambridge, MA 02139

Dr. Tjeerd Plomp
Twente University of Technology
Department of Education
P.O. Box 217
7500 AE ENSCHEDE
THE NETHERLANDS

Dr. Martha Polson
Department of Psychology
Campus Box 346
University of Colorado
Boulder, CO 80309

Dr. Peter Polson
University of Colorado
Department of Psychology
Boulder, CO 80309

Dr. Steven E. Poltrock
MCC
9430 Research Blvd.
Echelon Bldg #1
Austin, TX 78759-6509

Dr. Joseph Psotka
ATTN: PERI-IC
Army Research Institute
5001 Eisenhower Ave.
Alexandria, VA 22333

Lt. Jose Puente Ontanilla
C/Santisima Trinidad, 8, 4 E
28010 Madrid
SPAIN

Dr. Lynne Reder
Department of Psychology
Carnegie-Mellon University
Schenley Park
Pittsburgh, PA 15213

CDR Karen Reider
Naval School of Health Sciences
National Naval Medical Center
Bldg. 141
Washington, DC 20814

Dr. Fred Reif
Physics Department
University of California
Berkeley, CA 94720

Dr. Lauren Resnick
Learning R & D Center
University of Pittsburgh
3939 O'Hara Street
Pittsburgh, PA 15213

Dr. Jeff Richardson
Executive Director
Center for Applied AI
Campus Box 419
University of Colorado
Boulder, CO 80309

William Rizzo
Code 712 NAVTRAEQUIPCEN
Orlando, FL 32813

Dr. Andrew M. Rose
American Institutes
for Research
1055 Thomas Jefferson St., NW
Washington, DC 20007

Dr. Richard Sorensen
Navy Personnel R&D Center
San Diego, CA 92152

Dr. Kathryn T. Spoehr
Brown University
Department of Psychology
Providence, RI 02912

Dr. Robert Sternberg
Department of Psychology
Yale University
Box 11A, Yale Station
New Haven, CT 06520

Dr. Albert Stevens
Bolt Beranek & Newman, Inc.
10 Moulton St.
Cambridge, MA 02238

Dr. Paul J. Sticha
Senior Staff Scientist
Training Research Division
HumRRO
1100 S. Washington
Alexandria, VA 22314

Dr. Thomas Sticht
Navy Personnel R&D Center
San Diego, CA 92152

Dr. Martin M. Taylor
DCIEM
Box 2000
Downsview, Ontario
CANADA

Dr. David Thissen
Department of Psychology
University of Kansas
Lawrence, KS 66044

Dr. Douglas Towne
Behavioral Technology Labs
1845 S. Elena Ave.
Redondo Beach, CA 90277

Dr. Paul Twohig
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Headquarters, U. S. Marine Corps
Code MPI-20
Washington, DC 20380

Dr. Beth Warren
Bolt Beranek & Newman, Inc.
50 Moulton Street
Cambridge, MA 02138

Dr. Barbara White
Bolt Beranek & Newman, Inc.
10 Moulton Street
Cambridge, MA 02238

LCDR Cory deGroot Whitehead
Chief of Naval Operations
OP-112G1
Washington, DC 20370-2000

A. B. Winterhauer
Research Associate
Electronics Division
Denver Research Institute
University Park
Denver, CO 80208-0454

Dr. Robert A. Wisher
U.S. Army Institute for the
Behavioral and Social Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

Dr. Martin F. Wiskoff
Navy Personnel R & D Center
San Diego, CA 92152

Mr. John H. Wolfe
Navy Personnel R&D Center
San Diego, CA 92152

Dr. Wallace Wulfeck, III
Navy Personnel R&D Center
San Diego, CA 92152

Dr. Joe Yasatuke
AFHRL/LRT
Lowry AFB, CO 80230

Dr. Joseph L. Young
Memory & Cognitive
Processes
National Science Foundation
Washington, DC 20550

Dr. Ernst Z. Rothkopf
AT&T Bell Laboratories
Room 2D-456
600 Mountain Avenue
Murray Hill, NJ 07974

Dr. William B. Rouse
Georgia Institute of Technology
School of Industrial & Systems
Engineering
Atlanta, GA 30332

Ms. Riitta Ruotsalainen
General Headquarters
Training Section
Military Psychology Office
PL 919
SF-00101 Helsinki 10, FINLAND

Dr. Michael J. Samet
Perceptronics, Inc
6271 Variel Avenue
Woodland Hills, CA 91364

Dr. Robert Sasmor
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Dr. Roger Schank
Yale University
Computer Science Department
P.O. Box 2158
New Haven, CT 06520

Dr. Janet Schofield
Learning R&D Center
University of Pittsburgh
Pittsburgh, PA 15260

Dr. Marc Sebrechts
Department of Psychology
Wesleyan University
Middletown, CT 06475

Dr. Judith Segal
Room 819F
NIE
1200 19th Street N.W.
Washington, DC 20208

Dr. Robert J. Seidel
US Army Research Institute
5001 Eisenhower Ave.
Alexandria, VA 22333

Dr. Ramsay W. Selden
Assessment Center
CCSSO
Suite 379
400 N. Capitol, NW
Washington, DC 20001

Dr. Sylvia A. S. Shafto
National Institute of Education
1200 19th Street
Mail Stop 1806
Washington, DC 20208

Dr. Lee Shulman
Stanford University
1040 Cathcart Way
Stanford, CA 94305

Dr. Zita M Simutis
Instructional Technology
Systems Area
ARI
5001 Eisenhower Avenue
Alexandria, VA 22333

Dr. H. Wallace Sinaiko
Manpower Research
and Advisory Services
Smithsonian Institution
801 North Pitt Street
Alexandria, VA 22314

Dr. Derek Sleeman
Stanford University
School of Education
Stanford, CA 94305

Dr. Edward E. Smith
Bolt Beranek & Newman, Inc.
50 Moulton Street
Cambridge, MA 02138

Dr. Elliot Soloway
Yale University
Computer Science Department
P.O. Box 2158
New Haven, CT 06520